27679

NATIONAL LIBRARY
OTTAWA

CANADA

BIBLIOTHÈQUE NATIONALE
OTTAWA

NAME OF AUTHOR........WILLIAM LOUIS LAI........

TITLE OF THESIS. THE EFFECTS OF DISCRIMINATION ACQUISITION WITH AND WITHOUT ERRORS ON REVERSAL AND NONREVERSAL SHIFTS IN PRESCHOOL AND SECOND-GRADE CHILDREN

UNIVERSITY..OF ALBERTA........

DEGREE FOR WHICH THESIS WAS PRESENTED.....PH. D........

YEAR THIS DEGREE GRANTED.........1976........

Permission is hereby granted to THE NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(Signed)..William Lai........

PERMANENT ADDRESS:

18211-80A ave.
Edmonton
alberta

DATED..March 4.......1976

NL-91 (10-68)

INFORMATION TO USERS

THIS DISSERTATION HAS BEEN
MICROFILMED EXACTLY AS RECEIVED

This copy was produced from a microfiche copy of the original document. The quality of the copy is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Canadian Theses Division
Cataloguing Branch
National Library of Canada
Ottawa, Canada K1A 0N4

AVIS AUX USAGERS

LA THESE A ETE MICROFILMEE
TELLE QUE NOUS L'AVONS RECUE

Cette copie a été faite à partir d'une microfiche du document original. La qualité de la copie dépend grandement de la qualité de la thèse soumise pour le microfimage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

NOTA BENE: La qualité d'impression de certaines pages peut laisser à désirer. Microfilmée telle que nous l'avons reçue.

Division des thèses canadiennes
Direction du catalogage
Bibliothèque nationale du Canada
Ottawa, Canada K1A 0N4

THE UNIVERSITY OF ALBERTA

The Effects of Discrimination Acquisition With and Without Errors on Reversal and Nonreversal Shifts in Preschool and Second-grade Children

by

William Louis Lai

©

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

EDMONTON, ALBERTA

SPRING, 1976

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Reasearch, for acceptance, a thesis entitled The Effects of Discrimination Acquisition With and Without Errors on Reversal and Nonreversal Shifts in Preschool and Second-grade Children submitted by William Louis Lai in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

....

Supervisor

....

....

T. Weckowicz
....

....

External Examiner

Date January 9, 1976

To Vivien and Stephen

Abstract

The purpose of the present research was to investigate the effects of errorless discrimination training on a concept-shift task in preschool and second-grade children. The technique of errorless learning was developed in operant research and has been immediately applied as an instructional tool to facilitate the learning of a variety of tasks in children. However, it also appears from the literature that this facilitation is limited to the acquisition of the current problem only and does not carry over to a subsequent related task. Insufficient attention has been directed toward the study of the specific transfer effects of errorless learning in children to permit a more precise evaluation of the technique. The present project was therefore designed to assess the transfer effects of errorless learning with a structured discrimination task that has received a considerable amount of theoretical and empirical attention.

It was hypothesized that dimensional learning would occur with errorless training, but that dimensional control resulting from errorless training would be stricter than that from conventional training. These hypotheses were tested by means of the reversal and nonreversal shift paradigm, on the basis that relative speeds of the reversal and nonreversal shifts may be used as an indicator of learning on a dimensional basis.

These effects were examined in two age-groups, preschoolers and second-graders. Two experiments with identical procedures were carried out, one on each of the age-groups. The procedure was essentially the standard two-choice simultaneous discrimination method. Errorless learning was achieved by gradually fading in the negative members of the stimulus pairs. Control subjects were trained in the conventional manner without fading. For all subjects, the shift problems were introduced immediately upon reaching criterion on the initial discrimination without fading.

The major findings were that, compared to conventionally trained subjects, errorless subjects in both age-groups had faster reversal than nonreversal shifts, and that nonreversal was much slower after errorless training than after conventional training. A further analysis of the response patterns to the individual stimulus pairs during shift suggested that there appeared to be a strong persistence, by the errorless groups, on the solution relevant for the initial discrimination, but no longer appropriate for the nonreversal phase. There was some suggestion in the data that girls, especially preschool girls, were more affected in this way by the errorless training than boys.

These findings were consistent with the hypothesis that errorless training tends to overstrengthen dimensional control. The results were interpreted as indicating that the errorless procedure used in this study had a function similar to other stimulus-enhancing procedures in highly emphasizing the initially critical features (dimensions) of the stimulus complex. Since these relevant stimulus dimensions were associated with reinforcement early in the fading training, the result was an extremely strong control of responding by these dimensions. This dimensional control was also very resistant to extinction. This interpretation could account for the results obtained. Thus, at shift, learning was relatively easy if the task was solvable on the same dimensional basis as the original discrimination (the reversal shift), but was extremely hampered if solution on a different dimension was required (the nonreversal shift).

The implication of these findings is that in spite of the initial benefits of the errorless technique, it may not be advantageous for the learner to do so because it does not seem to promote adaptive behavior according to changing environmental demands. Several limitations of the present study were noted and suggestions for further research into the relationship between errorless training and dimensional control were made.

Acknowledgements

I would like to thank the following people who have so graciously allowed me to recruit subjects under their charge and made room available for the experiments reported in this thesis: Mrs. Rita Ratti, Teacher, Springhill Nursery School; Dianne Daley, Director, Students' Union and Community Day Care Centre; Mrs. Debbie MacDonald, Downtown Day Care Centre; Mrs. Mary Hull, Director, Community Day Care Centre; Mrs. Anna Beres, Director, Jasper Place Day Care Centre; Mrs. Lynn Scanlon, Director, South Edmonton Child Care Centre; Mrs. Anne Pura, Supervisor of Business Education, Edmonton Separate School System; Mr. M. A. Niebrugge, Principal, St. Angela School; Mr. R. Messier, Principal, St. Mathews School.

I am indebted to Mr. Joseph Kisilevich and Mr. Pat Wong, technicians with the Psychology Shop, University of Alberta. Their ingenuity and assistance in the construction of the apparatus is gratefully appreciated. Thanks are also due to Mr. Berry Hsu, our departmental programmer, for his valuable advice on computer usage.

To Dr. B. K. Sinha, my supervisor, goes my heart-felt gratitude for a schedule of continuous encouragement and assistance, not only during the course of this research, but throughout my graduate-student days (years, as a matter of fact) at the University of Alberta. I would also like to

thank the other two members on my Supervisory Committee, Drs. T. E. Weckowicz and C. Beck for their thought-provoking comments and constructive criticisms.

TABLE OF CONTENTS

PAGE

INTRODUCTION 1

Traditional Conceptions of Discrimination Learning . 1

Errorless Discrimination 3

Errorless Technique Applied to Children 12

Discrimination Learning in Children 19

Errorless Learning and Discrimination Learning in Children: Rationale of the Present Study 25

Purpose of the Present Study 35

Hypotheses 38

METHOD 41

Apparatus 41

Stimulus Materials 43

Experimental Design 44

Subjects 44

Procedure 45

RESULTS 51

Experiment I 53

Experiment II 70

DISCUSSION 79

CONCLUDING REMARKS 93

REFERENCES 98

APPENDIX 1. TABLES 105

APPENDIX 2. FIGURES 135

APPENDIX 3. ARRANGEMENT OF STIMULUS PRESENTATIONS 146

LIST OF TABLES

Table	Description	Page
1	Summary of the Analysis of Variance on Preshift Trials to Criterion for Preschool Subjects, Transformed Data	106
2	Summary of the Analysis of Variance on Shift Trials to Criterion for Preschool Subjects, Transformed Data	107
3	Mean Trials (raw scores) to Shift Criterion of Preschoolers as a Function of Preshift Training and Shift Conditions	108
4	Mean Trials (raw scores) to Shift Criterion of Preschoolers as a Function of Sex, Preshift Training and Shift Conditions	108
5	Summary of the Analysis of Variance on Shift Trials to Criterion of Preschool Girls, Transformed Data	109
6	Summary of the Analysis of Variance on Shift Trials to Criterion for Preschool Boys, Transformed Data	109
7	Summary of the Analysis of Variance on Shift Errors to Criterion for Preschool Subjects, Transformed Data	110
8	Mean Trials (raw scores) to Shift Criterion of Preschoolers as a Function of Preshift Training and Shift Conditions	111
9	Mean Errors (raw scores) to Shift Criterion of Preschoolers as a Function of Sex, Preshift Training and Shift Conditions	111
10	Summary of the Analysis of Variance on the Proportions of Errors during Shift for Preschool Subjects	112
11	Summary of the Analysis of Variance on the Number of Perseverative Errors on the First Changed Shift Pair, Preschoolers, Transformed Data	113
12	Summary of the Analysis of Variance on Shift Trials to Criterion, Excluding Perseverative Errors on the First Changed Shift Pair, Preschoolers, Transformed Data	114

13 Summary of the Analysis of Variance on Shift Trials to Criterion, Excluding Perseverative Errors on the First Changed Shift Pair, Preschool Girls, Transformed Data 115

14 Summary of the Analysis of Variance on Shift Trials to Criterion, Excluding Perseverative Errors on the First Changed Shift Pair, Preschool Boys, Transformed Data 115

15 Mean Trials (raw scores) to Shift Criterion of Preschoolers as a Function of Sex, Preshift Training and Shift Conditions, Excluding Perseverative Errors on the First Changed Shift Pair 116

16 Number of Preschool Spontaneous Shifters as a Function of Preshift Training Methods 117

17 Number of Preschool Spontaneous Shifters as a Function of Original Relevant Dimensions 117

18 Number of Preschool Spontaneous Shifters as a Function of Preshift Training Methods and Relevant Dimensions 118

19 Summary of the Analysis of Variance on the Proportions Correct of the R versus NR-C Subproblems for Preschool Subjects 119

20 Summary of the Analysis of Variance on the Proportions Correct of the R versus NR-U Subproblems for Preschool Subjects 120

21 Summary of the Analysis of Variance on the Proportions Correct of the NR-C versus NR-U Subproblems for Preschool Subjects 121

22 Average Subproblem Correct Proportions, Preschool Subjects 122

23 Summary of the Analysis of Variance on Preshift Trials to Criterion for Second-grade Subjects, Transformed Data 123

24 Summary of the Analysis of Variance on Shift Trials to Criterion for Second-grade Subjects, Transformed Data 124

25 Mean Trials (raw scores) to Shift Criterion of Second-graders as a Function of Preshift Training and Shift Conditions 125

26 Summary of the Analysis of Variance on Shift Errors to Criterion for Second-graders, Transformed Data 126

27 Mean Errors (raw scores) to Shift Criterion of Second-graders as a Function of Preshift and Shift Conditions 127

28 Summary of the Analysis of Variance on the Number of Perseverative Errors on the First Changed Shift Pair, Second-graders, Transformed Data 128

29 Summary of the Analysis of Variance on Shift Trials to Criterion, Excluding Perseverative Errors on the First Changed Shift Pair, Second-graders, Transformed Data 129

30 Number of Second-grade Spontaneous Shifters as a Function of Preshift Training Methods 130

31 Number of Second-grade Spontaneous Shifters as a Function of Preshift Training Methods and Relevant Dimensions 130

32 Summary of the Analysis of Variance on the Proportions Correct of the R versus NR-C Subproblems for Second-grade Subjects 131

33 Summary of the Analysis of Variance on the Proportions Correct of the R versus NR-U Subproblems for Second-grade Subjects 132

34 Summary of the Analysis of Variance on the Proportions Correct of the NR-C versus NR-U Subproblems for Second-grade Subjects 133

35 Average Subproblem Correct Proportions, Second-grade Subjects 134

LIST OF FIGURES

Figure Page

1 An example of the reversal and nonreversal shifts ..136

2 Mean trials (raw scores) to shift criterion of preschool children as a function of Training and Shift conditions 137

3 Mean trials (raw scores) to shift criterion of preschool children as a function of Training, Shift and Sex 138

4 Mean trials (raw scores) to shift criterion of preschool chidlren as a function of Training, Shift and Sex, excluding multiple errors on the first changed shift pair 139

5 Subproblem learning curves of preschool children ...140

6 Subproblem learning curves of preschoolers as a function of Training and Sex 141

7 Mean trials (raw scores) to shift criterion of second-grade children as a function of Training and Shift conditions 142

8 Mean trials (raw scores) to shift criterion of second-grade children as a function of Training and Shift conditions, excluding perseverative errors on the first changed postshift pair 143

9 Subproblem learning curves of second-grade children 144

10 Subproblem learning curves of second-graders as a function of Training and Sex 145

Introduction

Traditional Conceptions of Discrimination Learning

Traditional learning theory, as represented by those of Hull (1943) and Spence (1936), has conceptualized discrimination learning as a dual process of acquisition and extinction through differential reinforcement. According to the Hull-Spence system, discrimination learning, as measured by differential responding, involves the interaction of excitatory and inhibitory response tendencies. Thus, an organism learns to respond appropriately to the positive stimulus (S+) by being reinforced for it, while responses to the negative stimulus (S-) are suppressed or extinguished by nonreinforcement. In addition, the excitatory and inhibitory tendencies are assumed to generalize to other stimuli as a function of either physical similarity or acquired mediated similarity. Every reinforced response to S+ produces a response tendency to S- and every nonreinforced response to S- also generates some inhibition of responding to S+. These generalized tendencies, however, are assumed to be weaker than the direct reinforcement/nonreinforcement effects. The net effective response strength to a given stimulus, then, is a function of the algebraically summated interaction between the excitatory and inhibitory tendencies accruing to that particular stimulus. Eventual mastery of the discrimination task is accounted for by the resolution of these competing tendencies over a series of trials, so that there is a stronger net excitatory response tendency to S+ and a stronger net inhibitory response tendency to S-.

With the interaction of excitatory and inhibitory tendencies and generalizations thereof as integral parts of the theory, errors during the course of discrimination learning seem inevitable from this viewpoint. The generalized effect of nonreinforced responses to S- is a reduction of response tendency to S+. At the same time, though the organism is not reinforced for responding to S-, the existence of a generalized excitatory tendency following responses to S+ results in a generalized increment in the probability of responding to S-. Thus, it follows from the traditional learning theory of Hull and Spence that the occurrence of errors is a function of both nonreinforcement and generalized reinforcement. It is evident from this line of reasoning that errors are an inevitable, and indeed necessary, part of the discrimination learning process.

While the inevitability of errors follows from the conditioning-extinction theory of Hull and Spence, other learning theorists have emphasized the significance of errors. For example, Keller and Schoenfeld (1950) stressed the importance of errors and their nonreinforcement in the development of a response probability difference to the two stimuli to be discriminated. Their emphasis on errors was carried to the point of providing unusually long and frequent S- presentations thus allowing for the fullest benefit of the effects of nonreinforcement.

This emphasis upon the occurrence of errors was developed to its extreme in Harlow's (1950, 1959) proposal of error factors. It should be noted that though Harlow was working on the learning set phenomenon, he believed that his theoretical treatment applies to all learning and is most relevant in the context of discrimination learning. According to Harlow, the mechanism responsible for successful discrimination is not, strictly speaking, learning per se; it is, rather, the extinction or the suppression of incorrect responses. In Harlow's view, the learner brings with him to the learning situation the correct habits already available, but also a number of other competing tendencies which obscure the correct responses. Mastery of the discrimination, then, consists in the elimination of these interfering tendencies through extinction until appropriate stimulus control eventually emerges. Thus, it can be seen that errors have played quite a major role in the theorizing on discrimination learning by leading theorists in this area.

Errorless Discrimination

Recent developments with errorless discrimination procedures question the generality of the theory that discrimination learning is primarily a function of acquisition and extinction, and the generalization effects thereof. The operant research of Terrace (1963a, 1963b) has questioned the role of nonreinforced responding in the acquisition of a discrimination. His work will be reviewed below.

But it should be noted that, even prior to Terrace's systematic study of this phenomenon, related ideas and procedures were utilized by researchers in the field. For example, William James (1890) found that he could obtain smaller two-point limens if training began with a widely separated pair of points and the distance between them was gradually reduced. Pavlov (1927) found that his dogs could discriminate much faster between two very similar circles by starting conditioning with the presentation of the positive circle and an ellipse, and then progressively reducing the difference between the circle and the ellipse, than by starting the conditioning with the final pair of similar circles. Schlosberg and Solomon (1943) and Lawrence (1952) found that rats could learn to discriminate two similar shades of grey with very few errors when the task was begun with black and white stimuli with a gradual change to the final pair of greys. Skinner (1938) obtained errorless performance by initiating the discrimination learning immediately after the response had been conditioned in the presence of the S+.

Terrace (1963a) attempted to identify the determinants of errorless discrimination in pigeons. Key-pecking was the response measure used and the stimuli to be discriminated were the colors red (S+) and green (S-)Terrace compared four methods of introducing the S-:

(1) Early-progressive -- approximately 30 seconds after the key peck had been conditioned, S- was gradually faded in, progressively changing from a dark key of 5 sec. duration to a fully bright green key of 3 min. duration.

(2) Early-constant -- discrimination training also began after the establishment of the key peck response in 30 seconds, but the duration and brightness of the S- started at their maximum values of 3 min. and full brightness respectively.

(3) Late-progressive -- S- progressively changing from a dark key of 5 sec. duration to full 3 min. illumination, but this fading was started only after 21 sessions of key peck conditioning to S+.

(4) Late-constant -- S- was introduced at its full duration and brightness after 21 sessions of key peck conditioning to S+.

In terms of S- responding (errors), the birds of the early-progressive group acquired the discrimination with virtually no responses to S-. The late-constant birds made the most errors. Between these extremes fell the early-constant and the late-progressive groups. Thus, the establishment of errorless discrimination was found to be most effective when fading was started early during training and proceeded gradually.

In a second experiment, Terrace (1963b) first trained pigeons on a red/green discrimination errorlessly, and then transferred the discrimination to one between a vertical and a horizontal line. The transfer was introduced in one of three ways: (1) Superimposition-and-fading -- the lines were superimposed upon the colors, vertical on red and horizontal on green, with the colors slowly faded out; (2) Superimposition only -- superimposition of lines on colors as above, but the colors were withdrawn suddenly without fading after a number of sessions; (3) Abrupt -- no superimposition, the change from colors to lines were abrupt. Upon reaching criterion, all birds in the above three groups were abruptly returned to the red/green discrimination.

The result of this experiment was that the birds which had superimposition-and-fading training acquired the vertical/horizontal discrimination with no errors, while the other groups made many errors. Furthermore, those subjects which learned the vertical/horizontal discrimination with errors all made errors upon returning to the red/green discrimination. The superimposition-and-fading group performed perfectly when returned to the red/green discrimination.

Results from these two experiments pose problems for the traditional generalization theory as reviewed above. According to the latter, there should be greater response generalization with an increased number of reinforced responses to the stimulus, and an increased similarity between the stimuli. In Terrace's first experiment described above, as training proceeded in the early-progressive group, the birds had a long history of reinforced responding to S+, and the physical similarity between S+ and S- increased. Therefore, one would predict from classical generalization theory that the early-progressive birds would start emitting responses to S- after a certain number of responses to S+ were reinforced or after the S+/S- difference had been reduced to a certain value. But this did not occur. Similarly, in the second experiment, classical generalization theory would predict that some responding to S- should occur once the colored backgrounds were faded out. However, this was not the case.

Terrace (1963a) suggested that the absence of responding to S- with the progressive fading methods produced a permanent "short-circuiting" of the S- extinction process. He attributed this short-circuiting to two factors. One is the absence of intermittent reinforcement resulting from responding to S-, which ordinarily would increase the resistence to extinction. The second factor is the superstitious conditioning of not responding to S-, explained as follows. Typically, the initial responses of the early-progressive birds to the dark key was jerking the head back away from the key. Since the duration of the first S- was short (2-5 sec.), S+ appeared soon after the pigeon abruptly pulled away from the key. The reappearance of S+ could have served as a conditioned reinforcer for the movement away from the key during S-, thus preventing S-responding.

Terrace noted a further problem for classical generalization theory. He maintained that inhibitory gradients cannot be used to explain the absence of responding to S- in the acquisition of both the color and the line discriminations.

Two other observations in these experiments had led Terrace (1963a) to state that learning with errors would result in "permanently faulty discrimination performance". In the first experiment, post-criterion error bursts were observed in birds not trained with the early-progressive procedure. In the second experiment, birds which made errors in the line discrimination phase also made errors when returned to the color discrimination which had been learned errorlessly prior to the line discrimination. Thus, Terrace concluded that "once errors occur during the formation of a discrimination, subsequent performance is permanently affected on that, and also on related discriminations" (Terrace, 1963b, p. 231). Implicit in this statement is that traditional conceptions of discrimination would lead one to make use of procedures which produce "permanently faulty discrimination performance".

Terrace's work has also revealed differences in stimulus control following discrimination learning with and without errors. He has attempted to show that S- acquires aversive properties when subjects experienced errors in training, but remains neutral under errorless learning procedures.

The first indication of the aversiveness of S- was the occurrence of "emotional" responses such as wing flapping and turning away from the key following discrimination with errors. During errorless learning, the usual behavior of the pigeons in the presence of S- was usually a slow settling down under the response key and quietly waiting for the next appearance of S+ (Terrace, 1963a).

Another phenomenon noted by Terrace (1963a, 1966) in his errorless birds was the absence of behavioral contrast. Earlier it had been demonstrated that a consequence of discrimination learning is an increase in the strength of the response to S+, in the form of either an increased response rate (Reynolds, 1961) or a decreased latency (Jenkins, 1961), accompanied by a decrease in the S- response strength. Neither of these effects occurred with Terrace's errorless birds, but both were evident in the birds which made errors. Terrace suggested that since contrast might be a manifestation of emotional responses generated by the aversiveness of receiving no reinforcement for responding to S-, the absence of contrast with errorless training might be interpreted as demonstrating the neutral nature of S-.

That the S- in errorless discrimination is not an aversive stimulus was further demonstrated by an absence of the peak shift in a post-discrimination generalization gradient. Typically, the peak of a generalization gradient obtained after discrimination training in which responses to S- are extinguished is displaced away from S+, in a direction that is also away from S- (Hanson, 1961; Pierrel and Sherman, 1962). With errorless training, however, Terrace (1964) found that the generalization gradient peaks right at S+. Terrace suggested that the peak shift away from S- which had undergone extinction was a shift away from an aversive stimulus, and hence the absence of a peak shift following errorless discrimination showed that S- was functioning as a neutral stimulus.

To further show that S- becomes aversive in discriminations with error, but not in those without errors, Terrace (1971) showed that only pigeons trained with errors would learn a response to remove, or escape from, S-. In this experiment, a successive vertical/horizontal discrimination was trained on the right-hand key of a two-key apparatus. During S- each peck on the left-hand key, illuminated by a homogeneous white light, turned off S- for five seconds but had no effect on the stimulus presentation of the right key. This escape from S- behavior was observed only in those subjects who learned the discrimination with errors. In contrast, the errorless group emitted near-zero responses to the key that would remove S-. Thus, Terrace concluded from this finding that S- was neutral for the errorless birds but aversive for those with errors.

A further piece of evidence presented by Terrace to support the neutral status of S- following errorless discrimination was the flat generalization gradient along the S- continuum. Based on Jenkins' (1965) definition of an inhibitory stimulus as a stimulus which controls the tendency not to respond in a manner analogous to the way that an excitatory stimulus controls responding, Terrace (1966b) reasoned that just as the control of excitation is determined by varying S+ and observing to what extent the tendency to respond decreases as the distance between a test stimulus and S+ is increased, the control of inhibition may be determined by varying S- and observing to what extent the tendency not to respond increases as the distance between a test stimulus and S- increases. Thus, a flat gradient would signify no stimulus control, while a U-shaped gradient with minimum at S- would be evidence that S- controlled not-responding. Accordingly, Terrace (1966b, 1971) conducted a series of experiments in which he showed that the U-shaped gradient along the S- continuum appeared only for subjects who learned the discrimination with errors while errorless subjects had flat and near-zero gradients. Terrace concluded that the flat gradients from the errorless group reflected a uniform lack of excitation and that stimuli from the S- continuum was neutral in the sense that they had neither inhibitory nor excitatory properties.

Errorless Technique Applied to Children

The errorless technique developed by Terrace was found readily applicable to instructional situations, especially those involving severely disturbed children. The method proved to be valuable in that it helps to eliminate the negative and learning-inhibiting behaviors that are part of the disorder. It also makes some form of learning possible in otherwise untrainable subjects. For example, Sidman and Stoddard (1967) were able to train a group of retarded children to distinguish between a circle and an ellipse with the fading method. Seven out of ten learned the discrimination, compared with only one out of nine in another group of similarly retarded children trained with errors. A even more dramatic result was reported by Hewett, Mayhew, and Rabb (1967) who painstakingly taught three autistic children three words using a rather elaborate fading procedure. They noted, however, that it took one subject three months and another subject seven months to master the three-word vocabulary. Acker (1966) also found that both autistic and normal children acquired color and alphabet discriminations more quickly with errorless than with conventional differential reinforcement methods. Dehn (1971) also reported both an increased performance level and a reduction of undesirable negative behavior when severely autistic children were taught color discriminations with an errorless method.

Where errorless procedures have been used on normal subjects, usually young children, the effort has mainly been to demonstrate the plausibility and perhaps superiority of the errorless method in the acquisition of some discrimination problem over the conventional procedure. Terrace's (1963a) four methods of introducing S- were applied by Karraker and Doke (1970) to teach kindergarten children to discriminate the letters b and d. They found that although not all subjects in the early-progressive group learned errorlessly, this method did minimize the proportion of errors in training compared to the other three methods. Powers, Cheney, and Agostino (1970) attempted to replicate Terrace's early-progressive procedure on normal preschool children. The only modification was the addition of white noise to the presence of S-. The intensity of white noise, however, was gradually decreased as S- increased in intensity. Three subjects were trained with traditional discrimination techniques and two had the errorless treatment. The result was that no child in the traditional group made the discrimination while both subjects in the errorless condition were successful in making the discrimination with very few errors. Employing a matching-to-sample task on normal preschool children, Moore and Goldiamond (1964) showed that angular discrimination, usually difficult for such young subjects, could be established if the incorrect match was gradually faded in. Nonfading of the stimuli occassioned no learning in the same subjects. No transfer task, however, was instituted after this errorless acquisition.

Despite its potential in alleviating extremely severe cases of learning difficulty and in establishing ontogenetically difficult discriminations as suggested by the above reports, there is little evidence to show that errorless procedures facilitate the transfer of learning, or even permit the same degree of transfer for which conventional methods seem to provide. There is no systematic study of the effects of errorless acquisition on the performance of a subsequent task in children. There are, however, a few isolated studies whose findings may be suggestive.

By fading in the negative stimulus to train retardates the acquisition of a series of two-choice discrimination problems, Bricker, Heal, Bricker, Hayes, and Larsen (1969) hoped to establish the classical learning-set behavior in thier subjects. But unfortunately, though the fading method appeared to be the most effective training method in the study for individual problem acquisition, there was little or no transfer from the training problems to test problems which were similar to the training problems but were presented without fading.

Gollin and Savoy (1968) compared the performance on a conditional discrimination problem in two groups of normal children. Subjects were trained to respond to one of a pair of cues presented simultaneously on a common background. Upon reaching criterion on this original discrimination, the same pair of cues was then presented on a different common background and the positive and negative values of the cues were reversed. After criterion was reached on this second task, the test of conditional discrimination was administered. The cue-background combinations were presented in a random order. Success in the conditional discrimination task required the subject to respond to the cue-background combinations which were reinforced during the first two training discriminations. The negative cue-background combinations of the first two training tasks were faded in for one group of children. For the other group, both stimulus displays started out with full brightness. The conditional discrimination was conducted at full brightness for both groups. Interesting differences were found between the two groups. In the first two training discriminations, there were more subjects in the fading group than in the nonfading group who learned the tasks without errors. But it was the fading subjects who made significantly more errors during the conditional discrimination. Furthermore, 40 per cent of the nonfading subjects performed the conditional discrimination without any errors, compared with only four per cent of the fading group. These findings led the authors to suggest that the gradual introduction of the S- in the fading procedure might have confined the subjects' attention to specific properties of the S+. This would affect the subject's transfer performance in the following way:

> Should fading techniques tend to confine the subject's attentional responses to specific attributes of S+, the subject is likely to be at a disadvantage when a transfer task is introduced,

> since those stimulus properties which have been reinforced early in the fading procedure are likely to control response. If the control is strong then the subject may not attend to other stimulus properties as they become manifest and, therefore, he may not acquire the attentional responses necessary for efficient transfer. (Gollin and Savoy, 1968, p. 450).

Following the lead of the Gollin and Savoy study, Cheney and Stein (1974) investigated the effectiveness of three fading and two nonfading procedures on oddity learning in kindergarten children. The fading procedures included: (1) initially saturating the odd stimulus with a red color and then fading out the red color; or (2) initially saturating the nonodd stimuli with a red color and then fading out the red color; or (3) gradually increasing the illumination of the nonodd stimuli. One nonfading group received training on a simpler oddity problem prior to transfer to the more complex oddity problem which was also the basic problem for all groups. The other nonfading group received standard oddity training on this basic problem throughout. For all groups, the first fifteen trials were either fading trials for the fading groups, or training trials for the nonfading groups. The second fifteen trials were test trials when the problem was the same for all groups and the stimuli were of equal brightness. The main finding from this experiment was that as long as the fading was in progress, the subjects in all fading groups performed on a near perfect level, but that their performance deteriorated rather drastically during the test trials when the fading was over. In comparison, the nonfading groups did not have a similar dramatic performance level druing training, but they did not deteriorate during the test trials either. Cheney and Stein (1974) therefore suggested that one of the limitations of fading procedures might be the restriction of the subject's attention to those properties of the stimuli present during the earlier part of training so that the emerging properties of the stimuli, which are critical for transfer, are not adequately observed.

The utility of stimulus fading as a training technique for the attainment of concepts was questioned by Schwartz, Firestone, and Terry (1971) who did not find fading procedures particularly helpful in teaching elementary school children the concept of bilateral symmetry. Although this task seemed to be on a much higher cognitive level than the ones usually given to children employed in experiments, it nevertheless may indicate a limit for the fading technique.

Finally, there is one somewhat positive finding in regards to the effects of fading on nursery children's transposition. Cole, Dent, Eguchi, Fuji, and Johnson (1964) trained their subjects to respond to the larger of two stimulus squares. The smaller, negative stimulus was faded in, starting with a 2/10-inch line and increasing in 30 equal increments until it was completed on the last training trial. The children were divided into three groups: (1) those given immediate transposition; (2) those receiving 15 additional training trials with both stimuli complete, followed by the transposition test; and (3) those receiving 30 additional training trials with complete stimuli, followed by transposition. The results showed that subjects in all groups demonstrated a significant tendency to transpose. One drawback of this study is the absence of a control group who should be trained on the conventional, nonfading method so that comparison could be made. Another problem is that with transposition the same dimension is relevant in the original as in the transposition tasks. Therefore, although this study seemed to indicate transfer subsequent to fading training, its generality was by no means established.

The above review of current literature on errorless learning as applied to children suggests the following conclusions. Training without errors tends to be highly effective for individual problem acquisition with a variety of tasks. In most cases it is far superior to conventional methods in terms of bypassing learning difficulties. The main drawback, and a crucial one, of the technique seems to be the problem of transfer. There is very little evidence to indicate that errorless learning is valuable as training for adaptive behavior. It would seem from presently available evidence that errorless training does not allow for the abstraction of commonalities in the stimulus complex as a basis for responding. The gradual introduction of the S- may not provide the subject with sufficient comparative experience to permit efficient transfer which requires the establishment of a dimension schema to mediate the subject's observational techniques in a manner that insures more complete comprehension of the properties of the stimulus field. The skills trained by the errorless method then would be rather specific in terms of stimulus control.

Discrimination Learning in Children

The theoretical rationale as well as methodology for the study of discrimination learning in children came largely from traditional concepts emerging from animal studies. Thus, the conceptions reviewed in the beginning section of discrimination learning applied to the earlier work on children in this area. It was soon discovered that theories arising out of the animal literature were not adequate to handle human data, and subsequent developments have generated theories specifically designed for discrimination learning in children. Nevertheless, most of these later developments do have their historical roots in the Hull-Spence school of thought.

There are three major contemporary approaches to discrimination behavior in children. Two of these emerged from Spence's theory; the attention model of Zeaman and House (1963), and the mediation theory of Kendler and Kendler (1962). Both formulations extend Spence's theory to encompass selective responding to critical features of the stimulus situation. Both posit selective mediational mechanisms but assign different conceptual properties to them. A third account, proposed by Tighe and Tighe (1968), is differentiation theory which is based on Gibson's perceptual theory (Gibson, 1950, 1959).

The first theory to be described is the Zeaman and House attention theory. On the basis of Hull's (1930) earlier work on reinforcement, Spence (1936) had formulated a single-stage theory of discrimination learning as a nonselective and continuous process in which direct associations between stimuli and responses are gradually strengthened and weakened by reinforcement and nonreinforcement respectively. Later on, in order to deal with the problems of pattern discrimination (Ehrenfreund, 1948) and the attack by noncontinuity theorists such as Krechevsky (1938), Spence (1952, 1956) formulated a sequential two-unit S-R conception, in which the subject initially oriented his receptors toward that part of the environment containing the relevant stimulus and then responded to this cue. For Spence, however, this initial observing response is simply an instrumental response which is nonselective, the function being only to expose the subject to the stimuli. In the light of related works of others such as Wyckoff (1952), Lawrence (1950), Sokolov (1960), and Sutherland (1959), Zeaman and House (1963) extended Spence's theory of orienting response to include selective responding to the critical features of the stimulus situation. In the Zeaman and House model, mediating attention responses operate selectively, screening out sources of stimulation so that at any one time the subject is attending to only a single dimension (e.g. color), a dimension being defined as "broad classes of cues having a common discriminative property" (Zeaman and House, 1963, p. 168). Thus, this brand of attention theory accounts for discrimination learning in children by positing the acquisition of two successive responses, the first attentional, the second an instrumental-choice reaction, the probability of both being a function of the reinforcement schedule.

The Kendlers' version of mediation theory is an extension of Spence's (1952) nonselective S-R association to account for discrimination learning in rats. Evidence from studies on human subjects, however, had consistently contradicted predictions made from Spence's assumptions (e.g. Buss, 1953, 1956; Harrow and Friedman, 1958; H. H. Kendler and D'Amato, 1955). This led Kendler and Kendler (1962) to postulate a mediational S-R model consisting of a mediated representational response to the test stimuli. This mediating response in turn generates implicit cues leading to the overt choice response. At first, this mediation was described by the Kendlers as covert and verbal in nature; the subject presumably somehow labels the cues of the stimuli internally and responds accordingly. However, studies on the effects of verbalization and labelling of stimulus cues did not confirm the Kendlers' hypothesis (Wolff, 1967). In view of this, Kendler and Kendler (1968, 1970) modified their mediation theory to involve some form of "cognitive ability". The labelling aspect of mediation is minimized in importance. Instead, "common symbolic nediating resppnses" leading to concept formation and its utilization are stressed. Presumably, these symbolic responses are the subject's covert representation of the common properties of the stimuli. An important feature of the Kendlerian model is that it considers ontogenetic and phylogenetic differences in discriminative behavior explicitly. Higher organisms, as older children and adults, are assumed to discriminate in the mediating manner just described, while lower organisms, such as rats and young children, would function in the strict S-R fashion according to Spence's postulations. Thus, the formulation proposed by the Kendlers provides for two processes in discrimination learning. Which one of these processes applies is dependent upon the ability of the organism to organize or conceptualize the stimulus materials. This ability in turn may or may not be a function of the complexity of the stimuli in question.

The differentiation theory of Tighe and Tighe (1966, 1968) views the fundamental process in discrimination learning as one of differentiation of the stimulus array. Facilitation of discrimination is a function of the subject's increased sensitivity to and utilization of the differential properties of the stimuli. These stimulus properties are assumed to be the physical dimensions of the stimuli, such as color, form, height, etc. An efficient discrimination is characterized by dimensional control of choice behavior according to the distinguishing features (dimension) common to the stimulus array and appropriate for problem solution. Tighe and Tighe (1969, 1970) have argued that the subject's ability to isolate the relevant stimulus properties is a function of his ability to abstract the invariant stimulus attributes. Furthermore, they maintain that older children are more able to select and abstract these common attributes in the stimulus complex than younger children whose responses are primarily determined by unique object-reward associations.

In concluding this review of discrimination learning theories in children, several commonalities among these theories must be noted. There are controversies among these theoretical positions, to be sure, most of which centre around the inferred process of discrimination. But, as will be made clear, it is the shared characteristics of these theories that are most pertinent to this study.

Firstly, all the proponents of the three theories share the same methodology. All of them use the discrimination-shift paradigm. This methodology essentially involves reinforcement of responses to one set of stimuli and nonreinforcement to another set. The stimuli are presented in a pair-wise fashion and usually two pairs of stimuli form a problem. One member of each pair is designated as the correct stimulus requiring a specific response. All four stimulus elements are related to one another along one or more dimensions, so that the correct stimuli have certain features (dimensions) in common, while the incorrect elements have other features in common. Thus, the subject is confronted with stimuli varying in several dimensions (e.g. color and form) and each of these dimensions is exemplified in the stimuli by two or more cues (e.g. red, blue - color; square, triangle - form). In such tasks, the subject is typically required to "abstract" or identify these dimensions, i.e., the relevant one, and to discriminate among its cues by associating with each an appropriate response. Acquisition of the problem is indicated by a series of correct responses to the correct member of each stimulus pair (criterion). At this point and usually without forewarning to the subject, the stimulus-reward associations are changed, or "shifted". The stimuli as well as the dimensions and their cue values used in this shift phase may be identical with or different from those used in the original discrimination. The objective of the shift phase is to observe the effects of learning the original discrimination on the shift, or transfer, task. The main interest here is to see how and if the subject, having responded consistently to the relevant dimension, utilize this information to solve a new problem.

The common use of such paradigms reflects the shared concerns of all three theoretical positions reviewed above. All of them recognize that there simply is apt to be more in the learning environment than the subject can take in and store. Hence some form of processing the stimulus information is necessary. The subject may have to actively select, or attend to, the stimulus cues appropriate for the occasion, and to organize these relevant and distinguishing features in some way in order to solve the problem. When the subject has achieved this and has been responding consistently, his behavior can be described as being under dimensional control. Furthermore, all three theoretical positions are interested in how such acquisition affects subsequent learning behavior, hence the use of discrimination-shifts. The use of the shift task is crucial, not only as a means of demonstrating the various idiosyncracies in each of the theories. It also reflects the common concern of all three positions for the transferrability of what is learnt in one instance to another, which, in essence, is a concern for human adaptive behavior.

Errorless Learning and Discrimination Learning in Children: Rationale of the Present Study

The above review shows that the errorless techniques have emerged from operant research, while the theorizing on discrimination learning in children has its antecedants in the traditional conceptions of discrimination learning and has evolved into systems that take into consideration the relationship between components within a stimulus complex. It has become quite obvious that successful discrimination performance is dependent upon the child's ability to isolate the properties that differentiate the stimuli, and unless the child attends to and organizes the stimuli he will be unable to determine what these properties are and to use them effectively. The merits of the operant approach, as far as errroless learning is concerned, lies in the development of a seeemingly powerful training technology. However, the operant approach has not addressed itself to the study of the effects of such technology on human conceptualization behavior. For example, in the operant method, generalization effects are usually evaluated by the generalization gradient along the same dimensional continuum as the training stimulus. This seems to bear little resemblance to a situation confronted by a subject in a discrimination task in which he has to deal with several stimulus dimensions occuring simultaneously, and to select the appropriate cues for the solution of the problem. In such a discrimination task, effects of generalization, and often the processes of discrimination learning are indicated by the subject's performance on a subsequent transfer task where discrete discrimnanda are also used. Thus, while the contributions of operant research on the development of the technology is acknowledged, it has not established the fact as to how such technology affects human discrimination behavior. The main purpose of this study, therefore, is to investigate how this training technology affects conceptual behavior as indicated by discrimination shifts in children.

The theoretical rationale for the present study is based on the common properties of the three most prominant discrimination learning theories designed for children, as discussed above. It was highlighted during the review that all three positions view the superior form of discrimination as one involving the "abstraction" of stimulus dimensions.

Whether one chooses to call this process "attention", "mediation", or "differentiation", the observable behavioral effects are the same. That is, discrimination performance is judged by consistent responding according to some critical stimulus feature which is objectively identifiable. Implicit in all three theories is the idea that discrimination learning is a function of dimensional control which is analyzable in terms of the relationships between stimulus characteristics and response patterns. Thus, for example, if the subject responds consistently to all the square instances of a stimulus array regardless of the color of the stimuli, his responding can be described as being under the control of the form dimension of the stimuli.

The major question posed by this study is whether or not dimensional control occurs with errorless training, and if so, the extent to which such control affects subsequent behavior. This question is important because the errorless technique promises to be a powerful instructional or remedial tool with immediate applications. But as indicated earlier, the consequences of such training on further learning are unclear at the moment. Evidence emerging from applications of the errorless technique to training children seem to suggest specific stimulus control for the current task only. There is no bearing on whether the control is or can be dimensional in nature. Neither is there evidence demonstrating the transfer effects of the errorless procedure on relearning.

The discrimination learning theories for children have not dealt with errorless learning directly. However, their rationale can be applied to such a learning situation to make predictions regarding subject performance. For instance, since the critical element in discrimination learning agreed to by all theories is the isolation of relevant stimulus dimensions and since errorless training emphasizes the characteristics of the positive cue, learning should be greatly facilitated. In terms of attention theory, errorless learning would provide for the immediate and initial reinforcement of observing responses, or attention, early in the discrimination process. This early and rapid development of attention would lead to the circular relationship between attentive behavior and reinforcement, with a resulting increased instrumental performance level. Thus, the stress on relevant dimensions of these theories applied to errorless techniques predicts a much superior acquisition compared to non-errorless procedures, as well as in difficult tasks in which the critical features of the stimulus complex is initially highlighted. These particularly enhanced critical features would quickly come to control responding. This effect seemed to have been confirmed by the studies reviewed above where errorless procedures were shown to have much merit in facilitating discrimination acquisition and in bypassing learning difficulties.

But as noted above, what is less clear is the the effects of errorless training on subsequent transfer. In the formulations cited above, proficiency of transfer is predicted if the dimensional control acquired during original learning is sustained in the transfer task and if the new task is solvable on the same dimensional basis as the original one. Transfer tasks solvable on a different dimension would then be comparatively slow as the subject has to reorganize the stimulus features. It seems reasonable, therefore, to assume the encouragement of early and appropriate dimensional control with errorless training. But to what extent dimensional control thus shaped facilitate or hamper transfer remains to be investigated. Although Gollin and Savoy (1968), and Cheney and Stein (1974) speculated that errorless training might result in restricted attention, and hence strict dimensional control, there have been an absence of studies particularly designed to explore these specific effects with the appropriate paradigm. On the other hand, a conceptual analysis of current discrimination theories suggests that certain specific transfer effects should follow from the successful establishment of discriminative behavior with dimensional control. The present study was therefore devised to investigate this problem.

The present study makes use of the reversal-nonreversal shift paradigm. The assesment of subject performance is made by comparing the differential speeds with which reversal (R) and nonreversal(NR) shifts are achieved after training on an initial discrimination to criterion. Figure 1 illustrates an example of these two kinds of shifts. In the initial discrimination, the subject is confronted with two pairs of stimuli differing in two dimensions, form and color. In the example, the relevant dimension is form and the positive stimuli are the squares of each pair. After learning this discrimination to criterion, the subject is shifted to one of two new discriminations with the same stimuli. In a reversal shift, the subject must now respond to the triangles which were previously negative, although the relevant dimension is still form. In a nonreversal shift, the subject is required to shift his responses to the dimension that was previously irrelevant. In the example, responses to the red instances of the color dimension are reinforced in the NR shift. It is interesting to note that despite their differential preference for the inferred discrimination learning mechanisms, all three theoretical positions would generate comparable predictions about the relative speeds of R and NR solution. That is, if the subject has attended to the original relevant dimension (the Zeaman and House attention theory), or if the subject has acquired a set of relevant symbolic mediating responses (the Kendlers' mediation theory), or if the subject has isolated and abstracted the common relevant stimulus features (the Tighes' differentiation theory), then the prediction is for a faster R than NR shift performance. The reasoning of each position with reference to the R/NR shift paradigm can be further explained.

According to attention theory, because successful discrimination learning entails appropriate attention to the relevant dimension and because an R shift, which is intradimensional, provides positive transfer of the attentional response, the R shift would be executed faster than the NR shift. The NR shift, which is extradimensional, demands the extinction of the previous attentional response and the re-acquisition of a new one, hence requiring more learning.

According to mediation theory (T.S. Kendler, 1963; Kendler and Kendler, 1968), if the subject has acquired a set of relevant symbolic mediating responses representative of the relevant dimension, the R shift would be faster than NR shift because in a reversal, the initial dimension maintains its relevance and so does the mediating response. Only the overt response needs to be changed, and since the experimental situation provides only one alternative overt response, this change provides no difficulty. If, on the other hand, the subject learns the initial discrimination in a single unit S-R fashion, establishing only direct connections between the external stimulus and overt response, the prediction is for a more difficult R than NR. This is because R requires the replacement of a response that has previously been consistently reinforced with a response that has previously been consistently extinguished. For the NR shift, previous training has reinforced responses to the newly positive and newly negative stimuli equally often. Strengthening one of these associations does not require as much extinction of its competitor as in a reversal, and therefore the NR shift would be accomplished more easily.

Differentiation theory (L.S. Tighe, 1965; Tighe and Tighe, 1972) relates the ease of executing an R shift to the subject's ability to detect and utilize the distinguishing features of the stimuli. Thus, a faster R than NR is predicted for subjects who are able to discriminate on the basis of the relevant dimension since it remains relevant in a reversal and he has only to learn a relation between one aspect of this feature and reinforcement; while in an NR shift, he must first redirect his attention to and then re-differentiate the now-relevant dimension. On the other hand, the subject who learns the original discrimination on the basis of mere object-reward connections would find the NR shift easier since only one of these associations is changed in NR, while both are changed in R.

It is clear, then, that predictions regarding the relative ease of the R and NR shifts deduced from the three viewpoints are strikingly similar. The main debate among these proponents centres on the inferred mechanisms of learning. However, one does not have to subscribe to any of the inferred mechanisms in order to make use of the theories. It appears that the concept of dimensional control underscores all of the postulations, especially with regard to the R/NR shift paradigm. That is, if the relevant dimensions are initially emphasized or somehow highlighted during training, then these dimensions would quickly exert control, and consequently one would find a faster R than NR performance.

An additional advantage of the R/NR paradigm is that it is amenable to what has come to be known as subproblem analysis of discrimination shift learning (Tighe, Glick and Cole, 1971; Tighe and Tighe 1972). This advantage renders the paradigm particularly appropriate for evaluating the extent to which subjects accomplish their transfer on the basis of particular S-R associations learned during training versus their tendency to transfer on the basis of some conceptual formulation of the problem. With reference to Figure 1, note that for the R shift, the reward relations which obtain in the initial discrimination are reversed for both of the pairs of discriminanda (subproblems), whereas for the NR shift, the object-reward relations are unchanged for one pair of stimuli and reversed for the other pair. It is possible for the subject to treat the two stimulus pairs as either one problem or independent subproblems which happen to occur on alternative trials during the course of training. If the problem is initially learned as two subproblems, there is twice as much to learn for the reversal problem in transfer because the NR shift requires relearning of only one subproblem while the reversal requires the relearning of two.

Research in subproblem analysis has revealed certain characteristic response patterns for reversers and nonreversers (Tighe and Tighe, 1972). In cases where an NR shift occurs more rapidly than the R shift (turtles, pigeons, rats, and most four-year-old children), the subjects have learned the problem by learning the S-R connections that are correct for each pair of stimuli independently. Performance of the nonreversal, unchanged pair (NR-U) remains near or at 100% during the course of transfer while performance on the nonreversal changed (NR-C) and reversal (R) pairs is quite similar, beginning at or near zero and increases gradually to 100%.

In older children (about ten years old), where R was faster than NR, not only did performance on NR-U show a sharp decline at the beginning of transfer, but the R subproblems were learned more rapidly than the NR-C subproblem. Moreover, the majority of ten-year-olds showed "spontaneous reversal", i.e., reversing their choice on their first postshift exposure to NR-U after experiencing nonreward on NR-C.

This pattern of results indicates that subproblems were not treated independently by subjects whose R shift was faster than NR shift, but were treated independently by faster nonreversers. It further suggested that some sort of conceptual mediation was facilitating the transfer patterns of the former subjects. Thus, by observing these subproblem performance patterns, the present study could aim to localize more precisely the effects of errorless learning on transfer.

Purpose of the Present Study

In the light of these considerations, the following propositions will be investigated in this study. If errorless training leads to dimensional control, then the R shift should be executed faster than the NR shift. A nonmediating single-unit S-R mode of learning would result in the opposite effect, i.e., NR faster than R. Given that errorless training leads to dimensional control, the question is whether this control is restricted to the original relevant dimension only, so that when confronted with a situation requiring a solution on a different dimensional basis (the NR shift), the task becomes very difficult, more so than if training was done non-errorlessly. Thus, a restricted-dimensional-control hypothesis would predict a much slower NR shift speed than both an R shift after errorless preshift discrimination learning and an NR shift after non-errorless pretraining. On the other hand, if errorless training does not produce a stricter control than those prepared by conventional training, then there should be no difference between the respective shift performances of subjects trained either way. Furthermore, there is the possibility that since the errorless procedure draws the subject's attention solely to the correct stimulus without regard for the negative stimulus, this procedure might not allow too much opportunity for the subject to develope a conceptual representation of the set of stimuli presented to him. It would, then, be quite possible for errorless procedures to result in very specific S-R learning in much the same way that Kendler and Kendler (1962) had proposed for younger children. If this is the case, then errorlessly trained subjects would be expected to execute an NR shift faster than an R shift. Thus, by comparing the shift performances of subjects trained on an errorless procedure, and by comparing these to those of subjects trained on a conventional method, it is hoped that some of the specific transfer effects of errorless learning in children could be observed.

In addition, it also appears worthwhile to investigate if there are developmental differences. Although attention theory does not postulate differences due to age as such, there are suggestions that either preschoolers do not mediate (attention or otherwise) or conceptualize at all (Tighe and Tighe, 1972), or that they mediate but acquire and extinguish their mediating responses slower than second-grade children (Campione, 1970; Dickerson, Wagner and Campione, 1970; Dickerson, Novik and Gould, 1972). Therefore, if errorless training does augment attentional mediating responses, the pursuant effects would be more prominent in preschool children. That is, for preschoolers the improvement of R performance after errorless training over R performance after conventional training should be greater than that of the second-graders. With second-graders, who may be expected to be able to conceptualize stimulus properties already without the help of the errorless technique, the effect of errorless training would be minimal, so that R performance is predicted to be equal in the two conditions. However, the assumption of strict dimensional control predicts that their post-errorless NR would be slower than if the original discrimination was learned conventionally. On the other hand, if errorless training should reduce learning to the specific mode, then it should retard the performance of older children by decreasing their relatively stronger tendency to mediate, so that their execution of the reversal would be slower and nonreversal faster than their conventionally trained counterparts.

A supplementary purpose of this study is an attempt to apply the recently published method of subproblem analysis of R/NR performance. It is hoped that this may provide further information about the manner of shift problem relearning underlying speed. The main interest here is in trying to discover whether errorless subjects would learn the discrimination on the basis of specific S-R connections or on the basis of a conceptual formulation of the problem.

In summary, the present study was designed to investigate the effects of errorless discrimination learning, in terms of dimensional control, on reversal and nonreversal shifts in two age-groups -- preschoolers and second-graders. Two experiments were planned. Experiment I involved preschoolers and Experiment II employed second-grade children. Both experiments had the same design and procedure. These will be described in the Method section below. The main hypotheses regarding the relationship between errorless learning and dimensional control were similar for both experiments. But specific predictions in terms of data following from these hypotheses would be slightly different for each of the experiments. This was explained earlier and will be further detailed below.

Hypotheses

The above discussion gave rise to two main hypotheses for this study. It was hypothesized that: (1) dimensional learning would occur with errorless training, but (2) that such dimensional control would be stricter than that produced by conventional training. These hypotheses were tested in terms of the relative speeds with which the R and NR shifts are executed. Speed is defined as the number of trials to criterion. These hypotheses are further elaborated for each of the experiments as follows:

Experiment I (preschoolers):

I-1. That dimensional learning would occur with errorless training.

This would be manifested by:

I-1a -- a faster R shift than NR shift after errorless training.

I-2. That dimensional control resulting from errorless training would be stricter than that from conventional training so that:

I-2a -- errorlessly trained subjects would execute an R shift faster than conventionally trained subjects; and

I-2b -- errorlessly trained subjects would execute an NR shift slower than conventionally trained subjects.

Experiment II (second-graders):

II-1. That dimensional learning would occur with errorless training.

As for Experiment I, the prediction was:

II-1a -- a faster R shift than NR shift after errorless training.

II-2. That dimensional control resulting from errorless training would be stricter than that from conventional training.

Because eight-year-olds can be expected to learn on a dimensional basis already, there may not be a greater degree of reversal facilitation by errorless training. Thus, the specific predictions following from this hypothesis are slightly different from those for Experiment I.

With reference to reversal, the specific prediction was:

II-2a -- that both errorlessly and conventionally trained subjects would execute the R shift at approximately equal speeds.

But the dimensional control resulting from errorless training would still be expected to be restrictive. This would be evident if:

II-2b -- errorlessly trained subjects would execute an NR shift slower than conventionally trained subjects.

No specific hypothesis was offered for the planned subproblem analyses since its use is exploratory in nature. The purpose is to see if subproblem patterns can disclose or clarify some of the less obvious reasons which may account for the relative transfer speeds. One would, however, expect that whenever a group of subjects tend to solve the R faster than NR, a pattern of nonindependent subproblem transfer will emerge, and that a pattern of independent subproblem learning will characterize subjects with a speedier NR shift.

Method

The apparatus, stimulus materials, and the standard experimental procedure were identical for both experiments. However, minor adjustments to the situational demands were made, which did not affect the uniformity of the procedure.

Apparatus

The apparatus was a grey wooden box measuring 24" x 17" x 14". Mounted on the front panel of the box were two Plexiglas windows measuring 3" x 3", two inches apart. Centred approximately three inches below the lower edge-line of the windows was a circular hole out of which marble-reinforcers were dispensed upon correct responses. A receptacle was attached immediately below the hole to catch dropping marbles. Behind the Plexiglas windows and on the back of the front wooden panel, rear-projection screens were tapped to allow for the slide projection of the stimuli. Microswitches were connected to the Plexiglas windows, activating a set of mechanisms when pressed. On a correct choice, the microswitches released a marble which would then fall into the receptacle. The microswitches were also wired to two light bulbs on a control box, permitting E to know which stimulus was chosen on any one trial. Pushing on either of the Plexiglas windows also terminated the stimulus presentation for that trial. This was accomplished by connecting the microswitches to a shutter placed immediately in front of the slide projector lens. On the control box, which was connected to the various mechanisms behind the front panel, were also a toggle switch and a push-button switch. The toggle switch allowed E to determine which of the stimuli would be correct according to a prearranged stimulus presentation list. The push-button switch opened the shutter for a new trial. Inside the wooden box and on one of the inner lengthwise walls was mounted a solenoid-operated gate housed in a metal box. The gate regulated the release of one marble per correct response. Incorrect choices did not activate the gate. A rubber tubing approximately 12" long ran from the gate-box to the centre hole delivering the marble. About 14" from the front edge and on top of the box, marbles were fed into the gate-box via a hole immediately above the gate-box. Marbles, were stored in a Plexiglas casing sitting on the top of the wooden box. The marble storage had a bottom hole which coincided with the hole leading into the gate-box, ensuring a continuous supply of marbles. The rear end of the wooden box was openable with a door which also served as a screen so that E could observe and record responses and manipulate stimulus presentation out of the subject's sight. The stimuli were projected onto the screens through this back opening by means of a Kodak Carousal AV800 slide projector, seated approximately seven inches beyond the edge of the rear opening.

Stimulus Materials

The stimuli consisted of two geometric shapes of two colors: squares and triangles were used with red and green. Pairs of stimuli, as illustrated in Figure 1, were made into slides. Fading of the negative stimuli was prepared by photographing the pair of stimuli with layers of tracing paper covering the to-be-faded-in member. Steps of fading was achieved by photographing the stimuli with one layer of tracing paper less than the previous one. Twelve fading steps were instituted for each pair of the problem, so that the first fading step was photographed with twelve sheets of tracing paper laid on the negative stimuli and the last step one sheet. The nonfaded member of the fading sequence, as well as both of the members of the nonfaded stimuli were photographed with no tracing paper covering. Each member of each stimulus pair was faded, in accordance with a counterbalanced design. When projected in sequence, the faded stimuli gradually became more and more apparent. Since there were two pairs to the problem and there were twelve fading steps, fading was complete by the 24th presentation pair, so that beginning with the 25th presentation pair, both members of each pair were fully exposed. Projected on the screen, the square measured 1-3/4 in. x 1-3/4 in. and the triangle had a 2-in. base and a 1-3/4 in. height.

Experimental Design

Identical for both experiments, the experimental design was a complete 2 x 2 x 2 x 2 factorial. There were two types of Training during the initial discrimination, fading and nonfading, designated as the F and NF groups respectively; two relevant initial dimensions, form and color; two sexes of subjects; and two Shift conditions, reversal and nonreversal, designated as R and NR respectively. The abbreviations, F, NF, R, and NR will be used individually and conjunctively to identify the principal subgroups. For example, F-R will refer to the group who had training during initial discrimination and were subsequently given a reversal task.

Subjects

For Experiment I, 64 preschoolers were recruited from seven day care centres in Edmonton. Half were males and the other half females. The subjects were assigned randomly to the treatment conditions except for sex. That is, there were equal numbers of each sex in every treatment combination. The mean CA of the sample was 54.3 months with a range of 43 - 64 months. Not included in the 64 were thirteen subjects who failed to learn either the initial or shift discriminations as described below. Four other subjects were dropped due to either equipment failure or experimental error. Three more subjects refused to continue with the task and were also excluded from the analyses.

The 64 school-age subjects in Experiment II were obtained from two Separate Schools in Edmonton. Again, they were randomly assigned to treatment conditions, except for sex. The mean CA of the sample was 95.64 months with a range of 91 - 100 months. Most of the subjects were in the second grade at the time of the experiment. There were six nonlearners. Three others were dropped due to equipment failure. These latter nine subjects were not included in the 64.

Procedure

The experiments were carried out on location. For Experiment I, it was done in a spare room or some quiet corner within the day care centre building cordoned off for the experiment and free from interference by other activities of the centres. Thus, except for the physical difference, the other experimental conditions such as freedom from disturbance and apparatus set-up were uniform for all subjects. For Experiment II, both schools were able to spare an office, which guaranted privacy for the experiment.

The subjects were run individually. At the beginning of the experiment, each subject was shown an assortment of prizes which included candy bars, little toys and trinkets such as rubber balls, beads, whistles, etc. The subject was told that he could win a prize from the assortment at the completion of the "game" that he was about to play. Then the subject was seated comfortably in front of the apparatus and given the following instructions:

> "This is a game in which you will try to win as many marbles as you can. Here is a dish for you to put your marbles in. (E handed S a small container.)
>
> Now (name of child), look at these windows. (E pointed to the two windows on the panel.) The way we play this game is that there will be pictures showing in these windows. This is how you find out which picture is a winner. When we start the game you will press on one of the windows which you think has the winning picture in it, like this or this. (E pressed each of the windows to demonstrate.) Now, you try. (S was requested to press each of the windows once.) If you are right, that is, if you pick the winning picture, a marble will drop out of this hole, (E pointed to the hole) and you may put it with the rest of the marbles you have here. (E pointed to the container.) If you are wrong, no marbles will drop out. Each time the pictures appear, you will have one turn to press, that is, you choose only one picture and push it only once, and try to choose the picture that will give you the marble. If you try, you can win a marble every time you choose, and I want to see how soon you can find a marble every time. Remember, you can pick a prize over there when you have finished the game. O.K.? Are you ready? (E answered questions if any.) O.K., we'll start. (After the first correct choice, E said): That's right, or, good. (After the first incorrect choice, E said): No, that's wrong. Try again."

None of the subjects in Experiment II had problems understanding the instructions, though a few hesitated before responding to the first presentation, in which case E prompted by saying: "Now you can choose one of the pictures and push on it", or some such encouragement. With the younger children in Experiment I, there were a similar number of hesitations. A number of them were also rather inquisitive about and were distracted momentarily by the workings of the apparatus. When this occured, E dealt with the situation by assuring S a look into the apparatus but only when the "game" was over. During the course of Experiment I, some subjects also asked E explicitly for the correct stimulus and were then met with a game-like but noncommittal reply.

The stimuli were presented according to a modified Gellermann series (Fellows, 1967) which restricted the number of successive correct stimuli in one position to a maximum of three, and which presumably guarranteed randomness as well as elimination of position responding. The fading groups received the same sequence of stimulus presentation, the only difference being the gradual fading-in of the negative stimuli. The counterbalancing procedure was employed. For the initial discrimination, half of the Training x Shift groups had form and the other half had color as the relevant dimensions. These relevant-dimension subgroups were divided again so that an equal number of subjects had each of the two cues on the relevant dimension as positive. For the reversal shift, the previous positive cue became negative while the same dimension (color or form) was retained. For the nonreversal shift, an equal number of subjects in each of the relevant-dimension subgroups had each of the cues on the new dimension as positive. These experimental arrangements are explained in Appendix 3.

A correction procedure was used in which the next pair was advanced only after a correct response. The criterion was ten successive correct choices. The criterion for the fading was also ten successive correct responses, but after the completion of fading, i.e., when both positive and negative stimuli were of equal intensity. After criterion was achieved on the initial discrimination, half of the subjects were given an R shift and the other half the NR shift. The shift problems were introduced with no ostensible break between the initial and shift phases. During this shift phase, all groups had stimuli at uniform and full intensity. Criterion was again ten consecutive correct responses.

In addition to the basic procedure described above, a training procedure was instituted upon the first error after the second 12-presentation block to facilitate learning on the original problem. This procedure involved the following: E showed the subject each of the two pairs of stimuli and, without mentioning the color or form of the cues, E pointed to the positive member and said: "This is the winning picture. It always gives you the marble." Pointing to the negative stimulus, E said: "This is the losing picture. It doesn't give you the marble." This special training procedure was similar to that used in several other studies (e.g., House and Zeaman, 1962; Campione, 1970, 1971; Dickerson, 1966, 1967; and Dickerson, Wagner and Campione, 1970). This special training procedure has the advantage of substantially reducing the number of nonlearners, while at the same time the procedure does not seem to affect the relative performance patterns in an R/NR shift design (Campione, 1971). This tactic can be further justified with reference to Dickerson (1967) who indicated that a large proportion of kindergarten subjects eventually required special training to reach criterion on the original discrimination anyway. This method is deemed preferable to the alternative of dropping nonlearners, which may result in a biased sample selection. This special training procedure was not employed during the shift phase. Instead, when five or six incorrect responses occurred in any 12-presentation block, E would say: "Remember, I want you to pay attention to the two pictures and see if you can win a marble every time."

In the NF group of Experiment I, special training was required for only five of the subjects. Two other subjects did not reach criterion at 150 trials even with the special training and were dropped. Two more subjects were dropped for not reaching the postshift NR criterion at 150 postshift trials. In the F group in Experiment I, ten subjects required the special training. Five other subjects given the fading treatment did not reach criterion by the 48th presentation-pair and were dropped, because for fading subjects to take that many trials to reach criterion would amount to learning the original problem with errors similar to the control subjects, and hence including such subjects would defeat the purpose of this study. Also eliminated were four more subjects who successfully completed the fading phase but failed to reach the postshift criterion. All four were given the NR shift.

With the second-grade children in Experiment II, special training was given to nine subjects in the NF group, and ten subjects in the fading group. There was one NF nonlearner and one NF nonshifter, both excluded. Also excluded were two subjects who could not learn the initial discrimination with fading, and another two who could not learn the NR shift after successful fading.

Results

In accordance with the previously stated hypotheses, the purpose of the two experiments was to explore the transfer effects of discrimination acquisition without errors compared to that with errors. Speeds of the R and NR shifts, in terms of trials and errors to criterion, were the major variables, supplemented by subproblem analyses. Experiment I investigated these effects in preschool children and Experiment II in second-grade children.

The schema for the presentation of results will be the same for both experiments. Firstly, the data on the initial discrimination will be reported. This will attempt to show that the fading procedure did occasion errorless learning on the initial discrimination. It will be further shown that the main subgroups within each experiment had a similar number of preshift overt responses. Then, the transfer performance will be reported in two ways. Trials and errors to the postshift criterion will be used to indicate the speed with which the shift task was completed, followed by an analysis of the subproblem learning patterns of the various subgroups.

With the trials and errors data in both experiments, analyses of variance were performed on the square-root transformed ($\sqrt{x} + \sqrt{x+1}$) scores. This was because hetereogeneity of variance was found in all cases and a square-root transformation was deemed appropriate for this kind of data (Kirk, 1968) to better meet the assumptions underlying the analysis of variance. Means, however, will be computed from raw scores to facilitate the interpretation of data.

Experiment I

Initial Discrimination

1. Fading

Of the 32 children who were in the F condition, 20 completed preshift training without errors. The remaining twelve made errors ranging from one to four. This small number of errors could be considered minimal. There was an equal number of erring subjects who were to be given the R or NR shifts (six in each case). The 32 subjects given the conventional simultaneous discrimination had a much higher proportion of errors, an average of 19.20 per cent, compared to 4.26 per cent for the twelve F subjects who made errors. Thus, the fading method had largely succeeded in enabling initial acquisition with no or very few errors.

2. Trials to Preshift Criterion

The criterion for the initial discrimination was ten successive correct responses. The 2 x 2 x 2 x 2 (Training x Shift x Dimension x Sex) analysis of variance performed on the transformed trials to preshift criterion (summarized in Table 1) revealed two significant main effects, those due to Training ($F = 6.4459$, $df = 1/48$, $p<.05$) and to Dimension ($F = 11.7785$, $df = 1/48$, $p<.01$); and a significant Training x Dimension interaction ($F = 10.3866$, $df = 1/48$, $p<.01$). On the whole, learning with form relevant was faster than with color relevant (mean trials = 26.59 and 38.81 respectively, raw scores). With reference to the Training x Dimension effect, an examination of the cell means indicated that the interaction was mainly due to the different acquisition rates between the form-relevant and color-relevant discriminations in the NF condition. While the F subjects required means of 35.00 for form and 35.75 for color, the corresponding means for the NF subjects were 18.19 and 41.88 respectively ($t = 3.64$, $df = 30$, $p<.01$, transformed data). The significant Training effect indicated a slightly smaller number of trials required by the NF group to reach criterion than the F group. An inspection of the raw score means showed, however, that as groups the F subjects had only about five more trials than the NF ones (35.38 versus 30.03). This may have been due to the fact that a more or less fixed number of trials was required for successful fading. But inspite of the statistical significance, the actual mean difference was not very great.

These results suggested that the fading procedure did achieve the desired effect of errorless discrimination acquisition and that both the F and NF groups had equivalent exposure to the stimuli prior to shift. It must be noted that there was no Training x Shift effect. It seemed safe, therefore, to assume that as groups the subjects assigned to the critical postshift independent variables were fairly well equated in terms of preshift overt responses.

Transfer Performance

The main purpose of the following analyses is to observe the relative speeds with which the F and NF groups executed the R and NR shifts. Trials and errors to the postshift criterion are taken as indicators of speed. A secondary aim is to observe the patterns of subproblem transfer during the initial transfer trials.

1. Trials and Errors to Postshift Criterion

The criterion for the transfer phase, as for the initial discrimination, was ten consecutive correct responses. The 2 x 2 x 2 x 2 analysis of variance on the transformed trials to criterion is summarized in Table 2. Four effects were statistically significant, namely, Training, Shift, Training x Shift, and Training x Shift x Sex. The cell means relevant to these significant effects are contained in Tables 3 and 4. It will be seen that overall (Table 3), the subjects who were presented fading stimuli before shifting took longer to reach the transfer criterion than those who had fully exposed preshift presentation, and that on the whole, an R shift was executed faster than an NR shift. Table 3 also shows that while the R shift was executed at equal speeds whether or not the subjects had fading during the preshift condition, the initially fading subjects needed approximately 25 more trials than the NF subjects to achieve the NR criterion ($t = 2.45$, $df = 30$, $p < .01$). Furthermore, R was faster than NR in each of the fading and nonfading groups ($t = 7.26$, $df = 30$, $p<.01$, for the former; $t = 3.63$, $df = 30$, $p<.01$, for the latter). These results obtain when sex is collapsed as shown in Table 3 and graphed in Figure 2. However, when the sexes were separated, as suggested by the significant three-way interaction, it appeared that much of the Training x Shift interaction was due to the retarded performance of girls who had been given fading at first and then an NR task (Table 4). A separate 2 (Training) x 2 (Shift) analysis on the girls' data revealed a strong interaction effect ($F = 14.5765$, $df = 1/28$, $p<.01$, Table 5). The boys' data (Table 6) showed a significant Shift effect but did not have any interaction at all. The shift performances of the two sexes, in raw trials to criterion, are illustrated in Figure 3. Comparisons among the girls' means produced the following results. R was faster than NR with fading ($t = 8.28$, $df = 14$, $p<.01$). With conventional training, R also tended to be faster than NR, but the t obtained fell short of statistical significance ($t = 1.74$, $df = 14$, $p>.05$). Furthermore, compared to conventional training, fading resulted in the girls requiring fewer trials for reversal ($t = 1.80$, $df = 14$, $p<.05$), but more trials for nonreversal ($t = 3.37$, $df = 14$, $p<.01$).

The analyses on the errors to criterion produced similar results reflecting the analyses done on trials. Table 7 summarizes the analysis of variance on the transformed data, and Tables 8 and 9 present the relevant raw score means. It is interesting to note that whereas the fading technique succeeded in enabling the subjects to acquire the initial discrimination with few or no errors, it was these same subjects who had significantly more errors during shift. This was also true when the proportion of errors was taken as a dependent variable for analysis instead of the absolute numbers of errors ($F = 6.8830$, $df = 1/48$, $p<.05$, Table 10). In terms of proportions, however, there was no other significant effect.

It should be noted that the factor of Dimension did not exert any effect whatsoever in the postshift analyses. It appears that though on the whole the subjects were a little slow orienting to color at first, once criterion was achieved, the dimensional differences did not affect their postshift performance.

2. Further Analyses on the Postshift Trials to Criterion

The correction procedure was such that the next pair of stimuli was advanced only when a correct response occurred to the present pair, thus allowing multiple errors on any one pair. It was noticed during the experiment that quite a number of F subjects perseverated their erroneous response to the first postshift changed pair. It was therefore decided to do an analysis on the number of perseverative errors on this pair. For the purpose of this analysis, the Shift conditions were collapsed, resulting in a 2 x 2 x 2 (Training x Dimension x Sex) design. This procedure is justifiable on the grounds that the Shift factor was not an active treatment variable in terms of the dependent variable in question. The first changed pair was equivalent for all subjects regardless of their subsequent shift assignment, and the type of shift to be imposed was in no way perceivable by the subject at the first changed pair. The analysis, summarized in Table 11, showed a significant Training effect (F = 23.7842, df = 1/56, p<.01, transformed data). The F group averaged 7.84 and the NF group 2.78 perseverative errors (raw scores) on the first postshift changed pair. There was also a .05 significant effect for Dimension (F = 5.5068, df = 1/56). The preshift form-relevant subjects had 6.40 and color-relevant subjects 4.22 perseverative errors (raw scores). However, because this difference seemed minimal and confounded with the significant Training effect, and since there was no other significant postshift effect associated with Dimension in the above postshift analyses, this particular significant Dimension effect may best be considered a spurious one.

In view of the perseverative errors on the first changed postshift pair, the possibility of these erroneous trials accounting for the trials-to-criterion results was explored by re-analyzing the trials-to-criterion data with the multiple errors on the first postshift changed pair counted as only one error. Table 12 presents the analysis of variance summary. Comparing Tables 2 and 12, it can be seen that essentially the same results were obtained, except for the main Training effect. The Training x Shift, and the Training x Shift x Sex interactions were retained, however.

Because of the retention of the three-way interaction, separate analyses of variance for each sex were run on the postshift trials to criterion without the multiple errors. Table 13 summarizes the boys' data, and Table 14 the girls' data. Means in raw scores appear in Table 15 and are graphed in Figure 4. For the boys, only the main Shift effect was significant, with higher mean NR scores for both training conditions (F = 20.2566, df = 1/28, p<.01). With the girls, overall NR shift was also slower than overall R shift (F = 52.2035, df = 1/28, p<.01). But the Training x Shift interaction was also significant (F = 18.0900, df = 1/28, p<.01). Further comparisons between the cell means of the girls' scores were therefore made using separate t tests, with the following results. With fading, R was very much faster than NR (t = 10.20, df = 14, p<.01). With nonfading, R also appeared to be faster than NR, but the magnitude of difference was not nearly as great as with the fading group, the t obtained (1.79) only barely reaching the .05 level of significance (with df = 14 and p<.05, t = 1.76). Further, F-R was faster than NF-R (t = 3.13, df = 14, p<.01), and F-NR was also slower than NF-NR (t = 3.09, df = 14, p<.01).

3. Summary of the Trials and Errors Analyses

The overall data confirmed prediction I-1a in that errorlessly trained subjects generally executed an R shift faster than an NR shift, thus supporting hypothesis I-1. The finding that NR performance after errorless learning was poorer than NR after conventional training was in accordance with prediction I-2b, supporting hypothesis I-2. However, the other prediction following from hypothesis I-2, that errorless subjects would have a faster R shift than conventional subjects (I-2a), did not receive support from the overall data. But sex differences were found. While the boys' data supported predictions I-1a and I-2b, the girls' data also supported predictions I-2a in that their reversal performance appeared to be facilitated by fading training, whereas the boys' was not. Thus, while the two major hypotheses received general confirmation, the degree of support appeared to differ with sex. Put in another way, it seems that the general effects of errorless learning were in the directions as predicted, but that the specific manifestations may vary with sex.

There were other findings obtained from the trials and errors data. Whereas fading succeeded in errorless discrimination training, it was the errorlessly trained subjects who had more errors during shift. Fading resulted in more perseverative errors as the reinforcement schedule was changed. But the main effects of fading regarding the relative speeds of shift problem solution, including sex differences, lasted beyond the extinction of perseveration.

Supplementary Analysis

The main objective of this analysis is to find out how the various groups of subjects treated the transfer subproblems. Specifically, it is intended to assess whether or not subproblems in the shift phase were learned independently as a function of the two preshift training methods. This may further elucidate the learning process underlying the trials to criterion analysis. Two types of analyses will be reported for this purpose, spontaneous shifting and graphic analyses of subproblem transfer.

a. Spontaneous Shifting

Independence of subproblem learning implies that experience of a change in reward contingencies on each subproblem is necessary to produce a change in response to that subproblem. To the extent that a change in correct response for one subproblem influences the subject's choice on the other subproblem, nonindependence is implied. Thus, spontaneous shifting may be regarded as an indicator of subproblem interdependence. A spontaneous shift, therefore, refers to a change in the choice for a specific subproblem on the first exposure to that subproblem during transfer, but after the correct alternative has been changed for the other subproblem. For the R shift, this amounts to examining Trial 2 of the postshift phase; if Trial 2 was correct on the first response, then the subject was classified as a spontaneous reverser. For the NR shift, the response to the first appearance of the unchanged pair (NR-U) after the first changed (NR-C) was examined; spontaneous reversers would respond incorrectly to this NR-U trial.

Of the total 64 subjects, slightly over half (35) of them shifted spontaneously. There were 17 spontaneous shifters from the F group and 18 from the NF group (Table 16). This difference was not significant. The spontaneous shifters were reclassified according to their original relevant dimension (Table 17), and dimension under the preshift training methods (Table 18). None of the chi-squares from these tables was significant. There were also no sex differences in spontaneous shifting (18 males and 17 females). Nor did the original training methods have differential effects on the sexes on spontaneous shifting.

b. Graphic Analysis of Subproblem Transfer

This part of the analysis attempts to ascertain the pattern of subproblem transfer during the first ten pair-trial presentations. This number of presentations was determined because after the tenth trial, subjects began achieving criterion precluding their inclusion in the analysis, and proportions thus obtained could be quite misleading. The graphs to be presented represent the proportion of subjects who responded correctly to the various subproblems on the trial-pair presentation number of that subproblem as shown on the abscissa. In the case of changed subproblems, the analysis begins with the first occurrence of the given type of subproblem during transfer. In the case of unchanged items, the first occurrence of such an item following a changed item represents Trial 1 on the graph. According to the method established by Tighe, Glick and Cole (1971), for the reversal curve, pair-trial refers to successive presentations of both stimulus pairs, every point on the curve thus involving an averaging over the two pairs. It requires that there be a sufficient number of subjects remaining at and before the twentieth presentation of the reversal trial-pair to enable a meaningful interpretation of proportions. In this study, since there was a substantial and rather rapid attrition of subjects beyond the tenth reversal trial-pair presentation as they reached criterion (keeping in mind that due to the correction procedure, presentations are not synonymous with trials used as a dependent variable for learning criterion), caution must be exercised in interpreting the end portions of the curves where the numbers of subjects represented were less than the full sample and unequal for some points on the curves. The last points on the curves where there was a full slate of subjects are marked with a vertical bar.

The visual interpretation of the graphic presentation of subproblem performance follows that of Tighe, Glick and Cole (1971); Tighe and Tighe (1972); and Cole (1973). These authors suggested certain features in the graphs which may characterize independent and nonindependent subproblem transfer. Independence is marked by a consistently high proportion of correct choices on the nonreversal unchanged pair (NR-U) from the beginning of shift, contrasted with a slow, gradual acquisition of correct responses to the changed pair (NR-C). In addition, performance on the reversal pairs (R) is very similar to that of NR-C. For subjects who learn the subproblems as one problem (subproblem nonindependence), the NR-U starts out with a low correct proportion, and performance suffers on both NR-U and NR-C. There is a necessary initial drop in R performance, but the R curve quickly surpasses the two NR curves.

Subproblem curves for the F and NF groups are depicted in Figure 5. Applying the preceding interpretation guidelines to the two panels of Figure 5, the following observations can be made. For both F and NF groups, reversal performance appears to be superior to both NR subproblems as indicated by the rapid acceleration of the R curves. Moreover, the R curves in each of the panels maintains its superiority throughout. This is consistent with the trials to criterion analysis, where it was found that the speed of reversal was faster than nonreversal for all subjects. The second feature to be noted is the relatively low starting NR-U correct proportion in both groups. These two features suggest nonindependent subproblem learning for both F and NF subjects. With reference to the comparisons between the two NR curves, it appears that the performance on the two NR subproblems was more equally disrupted in the NF group. With the F subjects, inspite of an improvement at Pair-trial 2, NR-C drops for the next three Pair-trials, while NR-U continued to improve. One additional feature of Figure 5a (the F group) is the relapse of NR-U at Pair-trial 6 down to the same level as on Pair-trial 1, just as NR-C reached beyond 60%.

Thus, the visual inspection of Figure 5 appears to suggest that there is little evidence of independent subproblem transfer in either the F or the NF group. The main feature distinguishing between the two profiles appears to be the manner in which the two NR curves separate in each of the two panels of Figure 5. It appears that for the F group, despite the improvement on Pair-trial 2, incorrect responses to the NR-C pair was still rather difficult to extinguish. While a higher proportion of correct responses to NR-U relative to that of NR-C may be an indication of subproblem independence, other features of Figure 5a suggest that other factors might be responsible for the separation between the two NR curves, especially when the separation occurs only in the middle portion of the graph. The rapid acceleration of the R curves, the wide difference in the relearning rates between the R and NR-C curves, and the low, initial correct proportion of NR-U are all characteristics of subproblem nonindependence.

Because there was an indication of a sex difference with the trials-to-criterion analysis, the subproblem learning functions were regrouped according to sex and preshift training, resulting in Figure 6. The differential effects on the sexes are quite apparent from Panels a and c of Figure 6. While the learning rates of the reversal problems of both males and females seem to be similar, the sexes differ more on the relearning of the NR subproblems. For the F males, there seemed to be little disruption of NR-U and NR-C relearning for the first five Pair-trials. With the F females, performance on NR-C dropped after Pair-trial 2 while that on NR-U was retained and improved. For both sexes, the NR-U curves declined rather sharply as the NR-C rate gradually caught up. With conventional training (Panels b and d), the subproblem pattern obtained for males bears a very close resemblance to that obtained by Tighe, Glick and Cole (1971) for ten-year-olds. That is, performance suffers on both NR-U and NR-C, with the R curve quickly surpassing both. These are assumed to be clear indications of subproblem nonindependence. With the NF females, there seems to be a higher degree of subproblem independence, with NR-U being consistently superior to NR-C. However, for at least the first five Trial-pairs, the shape and position of the R curve as well as the fact that the NR-U did not start with a high proportion correct indicate some degree of nonindependence also. Therefore, it appears that subproblem independence or nonindependence is a matter of degree. In this case, it appears that compared to boys, girls showed a higher degree of independent subproblem learning when conventionally trained than when errorlessly trained.

c. Statistical Analysis of Subproblem Transfer

The graphic analyses in Figures 5 and 6 provided a good visual representation of different patterns of subproblem learning. But in order to assess the reliability of the differences in learning rates among the subproblems, a measure of subproblem transfer for each kind of subproblem for each subject suitable for statistical analysis was obtained. The measure chosen, in accordance with the graphic representations, was the proportion of correct responses during the first ten subproblem Trial-pairs, or up to the last Trial-pair if the subject reached criterion prior. Thus, the NR subjects were assigned two scores -- a proportion correct for his five NR-U Trial-pairs, and one for his five NR-C Trial-pairs. An R subject received a single score -- the proportion correct for his twenty reversal pairs (the first ten trials on each pair). These raw scores were treated with an arcsin transformation deemed approproiate for proportions (Kirk, 1968).

Separate analyses were carried out in three parts. First, a comparison was made of performance on the R and NR-C subproblems. Then, R and NR-U were compared. Finally, the relative rates of NR-U and NR-C were examined. It will be recalled that subproblem nonindependence is characterized by the superiority of R over both NR-U and NR-C, and independence by the similarity between R and NR-C, and by the superiority of NR-U over NR-C. The analysis of variance design for the first two comparisons was the same -- a 2 x 2 x 2 factorial (Training x Sex x Subproblem). The design for the NR-U/NR-C analysis was a split-plot factorial with two between-subject factors, Training and Sex (each with two levels); and one within-subject factor, Subproblem.

Tables 19, 20, and 21 present the analyses of variance summaries, and Table 22 contains the raw cell means. The statistical analyses had significant main effects only. They showed that overall, R was superior to NR-C (F = 37.0677, df = 1/56, p<.01), and to NR-U (F = 11.0534, df = 1/56, p<.01). These results substantiate the previous visual interpretation that all subjects tended to learn the subproblems nonindependently. However, Table 22 also indicates that generally, NR-U was also superior to NR-C, although not to the same extent as the superiority of R over NR-C (F = 5.3200, df = 1/28, p<.05). This finding may have reflected the varying degrees of subproblem independence as explained earlier.

d. Summary of Subproblem Analyses

The subproblem analyses seemed to have elucidated the results obtained from the trials to criterion analyses to some extent. Thus, the generally faster R than NR performance by all groups may be due to the suggestion that the subjects on the whole tended to treat the subproblems as a unit rather than separately. With the number of trials, it was found that original training had little differential effect on the boys' NR shift, but fading prolonged the girls' NR relearning. The subproblem analyses suggested that though F and NF boys required approximately an equal number of trials to reach the NR criterion, their manner of doing so was somewhat different. The F boys demonstrated a rather rapid acquisition of both NR pairs. But this early relearning proved to be unstable as performances on both NR pairs dropped rather markedly after the fifth Trial-pair, and had to be reestablished again. The NF boys, on the other hand, had their learning on both NR pairs disrupted early during the transfer phase, followed by a gradual reacquisition. With the girls, it seems that there was a higher degree of subproblem independence in the NF group than in the F group. Both the patterns of Figure 6d and the cell means of Table 22 attests to this tendency. In fact, Table 22 shows that the F girls had the best R learning rate of all the subgroups. This may suggest that the reason for the retardation of the girls' NR acquisition may be related to or the same as that responsible for the facilitated reversal.

Experiment II

The subjects for this experiment were eight-year-old second-graders. The principles and methods of data analysis were the same as those of Experiment I.

Initial Discrimination

1. Fading

Eight of the 32 subjects given the preshift fading training learned the initial discrimination with no errors. Twenty-four made errors ranging from one to four, with the majority making only one error (eleven subjects). There were an equal number (12) of erroring subjects who were to be assigned to the post shift R or NR conditions. Fifteen of the subjects who made errors were in the preshift form-relevant and nine in the color-relevant groups. The average error rate of these 24 subjects amounted to only 5.06 per cent of their total responses, while the error rate for the NF group was 22.11 per cent.

2. Trials to Preshift Criterion

As with the previous experiment, the criterion was ten successive correct choices. A 2 x 2 x 2 x 2 (Training x Shift x Dimension x Sex) analysis of variance applied to the transformed data produced two significant effects; Dimension ($F = 4.3839$, $df = 1/48$, $p<.05$), and Training x Dimension ($F = 5.6223$, $df = 1/48$, $p<.05$). This analysis is summarized in

Table 23. Generally, the form discrimination was learned faster than the color discrimination (mean trials = 32.5 and 40.97 respectively, raw scores). The significant Training x Diemnsion interaction was mainly due to the differential acquisition rates between the two Dimension conditions of the NF group. While the F subjects required means of 35.69 trials to reach the form criterion and 36.56 trials to reach the color criterion, the corresponding means for the NF group were 28.31 and 46.38 respectively. This latter difference was significant (t = 2.24, df = 30, p<.05, transformed data). There were no other significant effects, including the Training x Shift interaction. This warranted the assumption of equivalent preshift response rates for subjects assinged to the critical shift conditions.

Transfer Performance

1. Trials and Errors to Postshift Criterion

The criterion was again ten successive correct responses. Table 24 summarizes the analysis of variance on the transformed trials to criterion. Two effects were significant; Shift (F = 41.5816, df = 1/48, p<.01), and Training x Shift (F = 6.3853, df = 1/48, p<.05). Table 25 presents the relevant cell means which are graphed in Figure 7. It can be seen here that an NR shift was generally more difficult to complete than an R shift. But the level of difficulty with the NR shift was substantially higher with fading training than with nonfading prior to shift (t =

2.74, df = 30, p<.01). Fading and nonfading seemed to have no appreciable differential effect on the trials to the R criterion. Further, in each of the Training groups, R was more rapid than NR (t = 2.09, df = 30, p<.05, for the F group; and t = 2.57, df = 30, p<.01, for the NF group). There were no other significant effects in Table 24, including sex differences which were found with the preschoolers.

The analysis on errors to postshift criterion is summarized in Table 26. Only the Shift effect was significant (F = 30.1426, df = 1/48, p<.01). An overall slower NR performance due to errors was indicated. Although not statistically significant, the Training main effect showed the same trend as the one for trials to criterion, as illustrated by the cell means in Table 27. Thus, the trend here is similar to the one found with the four-year-olds, i.e., fading the pretraining negative stimuli tended to produce a larger number of postshift errors.

As with the preschoolers, Dimension did not have any effect in the above analyses on postshift data. Similar to the preschoolers, the NF eight-year-olds also found the initial color discrimination somewhat more difficult, but once they had learned the discrimination, dimensional differences as such did not exert an effect on their postshift performance.

2. Further Analysis of Trials to Postshift Criterion

Errors on the first changed postshift stimulus pair were also analyzed with the 2 x 2 x 2 design, collapsing the Shift conditions. It was found that prior stimulus fading produced significantly more perseverative errors on the first changed pair, averaging 3.59 errors, versus 1.81 for the NF group ($F = 5.43$, $df = 1/56$, $p<.05$, Table 28). There were no other significant effects.

The trials to criterion were then re-analyzed with multiple errors on the first changed pair counted as one. Table 29 and Figure 8 contain these results. A comparison between Tables 29 and 24 revealed the same significant effects, suggesting that the results from Table 24 were not due to the increased number of trials on just the first changed postshift pair, beyond which the treatment effects were still sustained.

3. Summary of the Trials and Errors Analyses

All the specific predictions with reference to shift learning speeds were confirmed by the results. Thus, both major hypotheses were supported. Unlike the results from Experiment I, no sex differences were found in terms of the trials and errors to postshift criterion. Similar to Experiment I, fading again tended to produce more postshift errors on the first changed postshift pair as well as throughout the course of the shift tasks.

Supplementary Analysis

This part of the analysis deals with the patterns of subproblem learning.

a. Spontaneous Shifting

Forty-six (72%) of the 64 subjects reversed spontaneously. This relatively large proportion of school-age spontaneous reversers is consistent with their overall faster R shift and with findings obtained by Tighe and Tighe (1972). Of the 46, 21 had fading experience and 25 received conventional training, as shown in Table 30. This proportion was not statistically significant as a chi-square performed on Table 30 resulted in a value of <1. The spontaneous reversers were also fairly evenly distributed with respect to the preshift fading and dimension assignments (Table 31, chi-square < 1, nonsignifcant).

b. Graphic Analysis of Subproblem Transfer

The same principles and methods of subproblem analysis as those used in Experiment I were applied to the data of the eight-year-olds.

Figure 9 presents the subproblem curves of the eight-year-olds. In both panels of Figure 9, there is no clear-cut evidence of subproblem independence. In both cases, NR-U started off with a low proportion correct, and R acquisition was rapid and superior to both NR curves. The eight-year-old

F subjects had subproblem learning functions (Figure 9a) not unlike those of their younger counterparts. That is, the acceleration of the R curves was rapid and superior to both NR curves; and a relapse of NR-C low rates after some improvement, contrasted with a continued improvement of the NR-U rate. Furthermore, as with the F preschoolers, the NR-U curve declined as NR-C rose.

With the conventionally trained school children, their subproblem learning functions (Figure 9b) showed a more consistent nonindependence pattern. Not only did the NR-U subproblem suffer a deep depression at the start of transfer, but the R subproblems were learned more rapidly than the NR-C subproblem. There is still a discrepancy between the NR subproblems with more disruptions on the changed pair. But these two curves are relatively close to each other and are not as dramatically at variance with each other at the mid-portion as those of the F subjects.

In summary the graphic subproblem analysis revealed some transfer processes which were not immediately obvious from the trials or errors to criterion data. While fading had no differential effect on R subproblem acquisition, which was indicated by the trials and errors data, the manner of NR subproblem acquisition was different. Neither training method seemed to have produced independent subproblem learning. But Figure 9 did indicate different disruptive effects on NR-C acquisition due to the difference in training methods.

c. Further Graphic Presentation

There were no differnces due to sex in the analyses on trials and errors for this experiment. But since the analyses on the data of the preschoolers suggested sex differnces, it was decided to replot the subproblem curves according to sex for the eight-year-olds as well, to see if any sex differences unobservable from the trials and errors data could be discovered. The main purpose was to find out if the school girls were as much affected by the fading treatment as the younger girls were. Figure 10 presents these curves. Overall, the figure indicates little evidence of subproblem independence for all groups. The F male second-graders seemed to have more problems with the re-acquisition of the NR-C pair than the preschool boys. But the most striking feature of Figure 10 focusses on Panel d where it can be seen that the female second-graders were much more affected by the fading procedure than males, in much the same way as the younger girls were. This finding was obscured in the trials and errors analyses.

d. Statistical Analysis of Subproblem Transfer

In accordance with the graphic presentations, the proportions of correct responses during the first ten NR pairs and first twenty R pairs were analyzed. Separate analyses were done on the comparisons between R and NR-C, between R and NR-U, and between NR-U and NR-C. Although sex differences did not show up in the trials-to-criterion analysis, the subproblem curves of Figure 10 seemed to imply differential effects of fading on the sexes in their NR learning. Therefore, the design for the analyses on these subproblem learning rates included sex as a factor. The analyses were run with an arcsin transformation of the raw proportions.

Tables 32, 33, and 34 present the summaries of the analyses of variance, and Table 35 contains the relevant means in raw scores. Only significant main effects were found with the analyses of variance. For all subjects, the R rate was superior to NR-C (F = 49.6651, df = 1/56, p<.01), and to NR-U (F = 11.1469, df = 1/56, p<.01). Overall, NR-U was also superior to NR-C (F = 16.7724, df = 1/56, p<.01). These findings support the interpretation that nonindependent subproblem learning occurred in all groups, with the peculiar feature of NR-U being superior to NR-C.

e. Summary of Subproblem Analyses

The findings obtained for the second-graders were very similar to those for the preschool children. Both the visual inspection of the subproblem curves and the statistical analyses showed that the R relearning rate was superior to both the NR subproblems. This confirms the trials-to-criterion data and suggests subproblem nonindependence. Nonindependence is further supported by the low initial correct response rate on NR-U. It appears that fading did not result in the subjects learning the subproblems individually. The graphic presentations, however, suggested that there may be differential training method effects which were not evident from any statistical analyses. The most striking feature of the subproblem patterns seemed to be the depressed mid-portion of the NR-C curve in contrast to the improvement of the NR-U curve. This particular effect is stronger in the girls' pattern than in the boys'. With the NF subjects, while the NR-U rate was also higher than the NR-C rate, the mid-portion contrast between these curves was not nearly as great as that of the F subjects. Thus, while the trials-to-criterion analysis found no reversal performance difference with respect to training methods, but a slower NR learning after fading than after nonfading, the subproblem analysis may suggest possible reasons for this latter difference.

Discussion

The major purpose of the present study was to investigate the effects of errorless discrimination acquisition on reversal and nonreversal shifts in children of two age groups. The postshift data therefore are pertinent to this major purpose. The results did seem to elaborate on the transfer effects of errorless learning. They supported some of the previously stated hypotheses and were inconsistent with some of the alternative ones. These findings will be discussed in more detail. But before doing so, certain aspects of the preshift data are also worth commenting on.

The fading procedure used in this research was generally successful in producing errorless discrimination learning in children. This provides one more instance demonstrating that, indeed, acquisition of a discrimination does not necessarily require an active extinction of incorrect instrumental responses. Though this study was not designed to test the Hull-Spence conditioning-extinction theory, the consistency of obtaining errorless discriminative behavior in this and other studies reviewed above do question the involvement of such a dual process under all learning conditions. Furthermore, this contention seems to hold not only for the Hull-Spence type of instrumental theory, but also for more cognitive models, such as attention or mediation theory, which also assumes a dual-process in the acquisition of relevant and extinction of irrelevant dimensional mediating responses. It appears that attention to the correct dimensions, and hence presumably the abstraction of common properties from the stimulus field, is possible even without an active extinction of attention to the irrelevant dimensions, at least for the type of task and subjects used here.

One implication of errorless learning with reference to dual process theories should be clarified, however. The fact that discrimination learning can occur without the extinction of S- responding, while pointing out certain weaknesses in the deductions made from these theories as exemplified by Terrace's work reviewed above, is not a direct refutation of dual-process theories, as implied by Terrace (1963a). The dual process of acquisition and extinction is designed for the situation where one response is reinforced while another is not reinforced. A refutation would have to involve differential reinforcement and not obtaining the projected excitatory and inhibitory tendencies. Since the errorless procedure does not actually entail differential reinforcement in the sense that neither responses to S- nor nonreinforcement occurs, the conditions under which dual-process theories are assumed to operate no longer hold. The success in producing discriminative behavior by eliminating responses to S-, therefore, cannot be construed to disclaim the function of a dual-process in other learning situations. There may simply be more than one way of learning a discrimination problem, depending on the occasion.

In any case, it appears that the advantage of the errorless procedure in the initial acquisition was confirmed by this study. The results of the two experiments showed that with the conventional training method, the preshift color discrimination was more difficult than the form discrimination for both age groups; whereas, with the fading procedure, both the color and the form discriminations were acquired equally well. This particular pattern of results is quite analogous to findings of other studies in which differential acquisition rates due to stimulus pretraining were examined. Earlier, it had been found (Lee, 1965; Suchman and Trabasso, 1966a) that preschool children had fairly definite dimensional preferences, generally preferring form to color after the age of four. Suchman and Trabasso (1966b) further found that normal kindergarten and nursery school children did in fact learn to discriminate stimuli better when these stimuli were on a preferred dimension than when they were on a nonpreferred dimension. Subsequently, Mumbauer and Odom (1967) found that, with preschool children, requiring subjects to verbalize the relevant but nonpreferred color dimension facilitated the initial color discrimination acquisition. Working with three-year-olds on height and brightness discrimination, Caron (1969) found that subjects trained on their nondominant dimension following sensitization to that dimension executed the initial discrimination much faster than subjects trained on the nondominant dimension without sensitization. The sensitization procedure consisted of having the children differentiate the cues of the to-be-relevant dimension by sorting the discriminanda prior to discrimination training. It appears, then, that the fading procedure employed in this study had a similar effect of orienting the children to the appropriate stimulus features, reducing dimensional differences and enabling them to learn the otherwise more difficult color discrimination with the same level of ease as the form discrimination.

The two basic questions posed for the present study were: (1) whether dimensional learning occurs with errorless training, and (2) if so, the extent to which this dimensional control is exerted on subsequent behavior in terms of discrimination shift performance. An attempt will be made to answer each one of these questions with the data provided by the present experiments. Other related observations will then be discussed.

With reference to the first question, the data suggest that dimensional control of responding did develop with fading training. Evidence for this assertion is provided by the rapidity of the R shift relative to the NR shift after errorless training. This was true for both age-groups. Generally, the post-errorless R performance was either equivalent or superior to post-conventional R. With the preschool girls, there seemed to be a further facilitation of such dimensional learning by fading as the F girls reached the R criterion somewhat quicker than the NF girls. Also in both experiments, fading did not decrease the tendency to reverse spontaneously as the numbers of spontaneous shifters in either training groups in both experiments were not significantly different. The subproblem learning analyses also indicated that there was no independent subproblem learning pattern associated with the fading training. In fact, there was a hint that the fading treatment might have helped the preschool girls to gain dimensional responding, as it was noted that without fading, the preschool girls seemed to have a higher degree of independent subproblem learning. This visual interpretation of the subproblem curves was further substantiated by the statistical analysis which showed that the reversal learning rate was much superior to the nonreversal rates.

As discussed in the Introduction, the superiority of the R shift over the NR shift can be attributed to dimensional control of responding. It appears that even without active experience with the negative members of the stimulus array, subjects were still able to discern the relevant stimulus features and respond accordingly. This does not support the earlier suspicion that errorless training might result in specific object-reward learning. Had this been the case, there would have been a much greater amount of interference in the R shift and a lesser amount of disruption in the NR shift following fading training. There would also be a pattern of subproblem independence.

The second question to be answered is how the dimensional control shaped by errorless training influences shift behavior. More specifically, the intent was to assess the strength of this control. The data of the present experiments suggest a very restricted control. The results discussed immediately above showed a certain degree of dimensional control. The severity of this control is indicated by the findings that generally NR proved to be more difficult when the original discrimination was learned errorlessly than when it was learned conventionally, and that the R/NR differential was much greater for the F groups than for the NF groups. With the exception of the preschool boys, the F subjects required about 50 per cent more trials to reach the NR criterion than the NF subjects. This conclusion is further supported by the subproblem analyses. The pattern of these functions suggest that the F groups did not learn the subproblems independently. Rather, there appears to be a very strong dimensional effect carried over from original training. The swift acceleration of the R curves is one indication of subproblem nonindependence. Another indication is the initial rates of correct responses to the NR-U pair. These latter rates were considerably less than perfect but higher than the NR-C rates. An independence pattern would show a perfect or near-perfect NR-U rate. Thus, the relative high rates of both correct responses to the NR-U pair and incorrect responses to the NR-C pair may signify that the subjects were still responding to the stimuli which were correct during the original discrimination. Moreover, the continued improvement of the NR-U rate combined with a deterioration of the NR-C rate during the mid-sections of the subproblem curves may further suggest the persistence of response control by previously appropriate dimensions. It must be noted that the NF groups in both experiments were also controlled by the original dimensions to some extent. But compared to the NF groups, the errorless groups took a much longer period of time to re-orient to the shifted dimensions. It appears from the subproblem curves that the errorless subjects persisted on the solution based on original learning for quite a while before it was disrupted. The new relevant dimension would then acquire control over responding at a later stage than in the case of the NF subjects who seemed to be able to extinguish their inappropriate shift responses much easier and earlier.

Interesting comparisons can be made between the findings of the present study and those reported in the discrimination shift literature. These comparisons may suggest reasons accounting for the transfer effects of errorless learning obtained here.

The present findings are quite similar to those obtained in the study of relative R/NR speeds as a function of procedures designed to augment the subject's reception of the stimuli in a variety of ways. Both of the studies cited earlier (Mumbauer and Odom, 1967; Caron, 1969) revealed that verbalizing or sensitizing to the nonpreferred or nondominant dimensions produced facilitated reversals. Tighe (1965) pretrained her five- and six-year-old children by requiring them to render same-different judgments on the training stimuli, thereby isolating for her subjects the to-be-relevant stimulus dimension. She found that children thus pretrained later executed a reversal shift more rapidly than those who were given irrelevant pretraining or who were presented with a nonreversal shift. Thus, the fading procedure used in the present experiments may have similarly served the function of increasing the salience of the appropriate stimulus cues for the subjects, resulting in faster reversals than nonreversals.

Another phenomenon in discrimination learning to which the present results bear close resemblance is the overlearning reversal effect (ORE). Numerous studies have demonstrated that overlearning the original discrimination facilitates reversal. In an exhaustive review and listing of the literature on human discrimination-shift performance, Wolff (1967) found substantial documentary support for the ORE in children. Two specific results are especially interesting. Eimas (1969) has shown that overtrained third-grade children, while learning the reversal faster than those not overtrained, also made more perseverative errors at shift. Shepp and Turrisi (1969) obtained data showing that with increasing amounts of overtraining, intradimensional shifts became easier while extradimensional shifts became more difficult, resulting in an increasing difference between the two shift conditions with the amounts of overtraining. Both Eimas, and Shepp and Turrisi favour the interpretation of their results by attention theory.

The fading method used did not facilitate reversal in the sense of reducing the number of trials to criterion compared to the conventional method, except in the case of the preschool girls. However, there are parallels between the ORE and the general findings from this study which are immediately obvious and rather compelling. There were more perseverative errors after fading. Postfading reversal was much faster than nonreversal, the latter being also much slower than NF-NR, so that the difference between the two shift conditions was greater for the F groups than for the NF groups. In the present study, the R shift following fading was faster than the R shift following nonfading in the four-year-old girls.

Of the three theories regarding discrimination learning in children, only two attempt to account for ORE explicitly. These are the attention and differentiation theories.

Attention theory explains the ORE in the following way (Fisher and Zeaman, 1973; Zeaman and House, 1963). It is assumed that the probability of the correct instrumental response reaches asymptote much faster than the probability of the correct observing response (attention), so that at criterion the former is usually very near asymptote while the latter is considerably below asymptote. Hence, overtraining is expected to produce a relatively larger increment of the observing response and negligible amounts of the instrumental response. From these assumptions, the prediction is for reversal learning after overtraining to be marked by a number of perseverative errors followed by acquisition of the new instrumental response. For the NR shift, there should be negative transfer of the previously relevant attentional response, and since this was made particularly strong by overtraining, impairment of the NR shift would be magnified. Without overtraining, reversal learning should be characterized by fewer perseverative errors, but by responses to irrelevant dimensions due to the weaker and hence more easily extinguishable attention to the previous relevant dimension at the start of shift. In essence, then, according to attention theory, overtraining has the effect of highly accentuating the subject's attention to the relevant dimension, and implicit in this rationale is the inference that overtraining actually produces very strict attentional control.

Differentiation theory (Tighe and Tighe, 1965, 1966) simply considers overtraining as one means of producing very strong control of responses by the original relevant dimensions. This would have the effects in facilitating reversal similar to prior pretraining in the differentiation of the stimulus features.

Though the Kendlers have not made use of their theory to account for the ORE directly, it was in fact partly due to the occurrence of the ORE that Kendler and Kendler (1968, p.220-221) modified their verbal hypothesis. From the Kendlers' new version of mediation theory, one would derive that overtraining would serve to increase the strength of the representational response which in turn would exert strict control over responding, hence reversal would be easier than nonreversal where the overstrengthened representational response would have to be extinguished before appropriate instrumental responses could occur.

To recapitulate, evidence from the present research supports the hypothesis that dimensional control can be developed via errorless training, and further suggests that errorlessly induced dimensional control is rather restrictive. There are certain similarities between the effects of errorless training and those of other variables on discrimination acquisition and transfer. These other variables have in common the function of emphasizing the relevant cue. It would appear that errorless training may be categorized with this type of variables. A theoretical analysis provided a possible interpretation of the present results. The reason for the strict dimensional control shaped by errorless training as represented by the fading procedure used in this study seemed to be related to the obvious possibility that fading has enormously emphasized the relevant dimensions, to the exclusion of other emerging properties of the stimulus array. The consequence of the procedure then would be a narrow as well as an extremely strong dimensional control of responding, which is also very resistent to extinction. Thus, at shift, learning was relatively easy if the task was solvable on the same dimensional basis as the original problem, but was extremely hampered if solution on a different dimension was required.

Thus far, interpretation was offered on the general results from the two experiments taken together. The similarities between the age-group results did enable a common set of interpretations. There were, however, differences between the two sets of data, as reported in the respective Results sections. There was, for instance, a sex diffference in the number of trials to shift criterion for the younger but not for the older subjects. There was a Training x Shift interaction for the four-year-old females, with the F-NR group taking much longer to reach criterion than the NF-NR group, while the F-R females reached criterion sooner than NF-R females (Figure 3). There was no such interaction for the male preschoolers, although males who had fading did have slower performances on both shifts than the nonfading males. It does not seem to be the case that dimensional control did not develop in these F boys, as the the quick acceleration of the R curve and the very low initial NR-U response rate are both indicators of subproblem nonindependence. It may be that, for some reason, the control was not as severe and resistant to extinction in boys as in girls. Visual inspection of the NF boys' subproblem curves (Figure 6a) did show that inspite of the early and steady relearning, the performance on the NR problems was rather unstable. From the subproblem curves in Figure 6 and in the light of the previous discussion on the effect of errorless learning on dimensional control, it appears that the preschool girls had a somewhat slower attention acquisition with traditional training, as evidenced by a borderline independent subproblem profile, but that they were much more affected by the fading training procedure so that their behavior was more rigidly controlled by dimensional responding than boys. In contrast, the four-year-old boys seemed to be less rigidly controlled although the lingering effects of fading can also be seen in the abrupt deterioration of NR performance at Pair-trial 6 in Figure 6a. It must be added here that although there were no sex differences in terms of the trials to criterion for the eight-year-olds, Figure 10 did indicate a more striking fading effect for the older girls. Thus, it appears that, for some as yet unknown reason, females were more susceptible to the fading procedure than males. These sex differences were quite unexpected. There is no a priori theoretical rationale for explaining these differences. Further investigations along this line should take this consideration into account for clarification.

Although the somewhat peculiar sex differences need to be investigated further, there is sufficient consistency in the present results to permit at least tentative conclusions. There seems to be little doubt that errorless training has its merits. This has been recognized in the literature review and further substantiated by the present study. The merits lie, however, mainly on acquisition, i.e., in teaching the subject the correct solution to a particular problem and in helping the individual to overcome certain learning difficulties. While the benefits of errorless training have been demonstrated, its transfer effects, as shown by the present study, are questionable. It seems that the advantages of errorless acquisition are outweighed by the disadvantages during transfer. The conceptual mediation acquired by the subjects with the errorless procedure so strictly controlled instrumental responses that a lengthy period of nonreinforcement was required to extinghish the mediational control. Hence, the behavior of the errorlessly trained subjects appeared to be rigid and inflexible. The errorless procedure does not seem to promote a generalized skill of orienting to the appropriate stimulus properties according to changing environmental demands. The present results, if substantiated, could therefore imply that inspite of the initial benefits of errorless learning, it may not be advantageous for the learner to do so because such training is not conducive to the development of adaptive behavior in new situations. Terrace's (1963a) contention that non-errorless discrimination acquisition would lead to permanently faulty discrimination is perhaps too strong an assertion.

Concluding Remarks

The present research represents only one effort to study the effects of errorless learning by means of one procedure and one paradigm, and as such its limitations must be recognized.

The first limitation to be considered has to do with the fading procedure as a means of establishing errorless learning. The fading method was used in this study because it closely resembles that employed by Terrace and others. But conceivably errorless learning could be achieved by other means than fading in the negative stimulus. One could, for instance, engineer the apparatus such that S- responding is rendered impossible, while at the same time present S- at its full values. With this arrangement, errorless learning may be achieved presumably without accentuating any of the stimulus features particularly. It would be interesting to find out how subjects would behave under these conditions, and hence to test the generality of the effects of errorless learning.

Another limitation is related to the paradigm. The use of the R/NR paradigm has both advantages and disadvantages. The main advantage is that this paradigm is a well-studied methodology within the area of discrimination learning in children. As noted in the Introduction, there is an abundance of theoretical background relating to R/NR shifts. Thus, specific predictions could be made and results obtained are readily interpretable. The paradigm also seemed to have an added advantage in the form of subproblem analysis, which promised to provide additional information about the learning process. But the R/NR paradigm is not without problems.

The main drawback of the R/NR shift paradigm is perhaps the fact that identical stimuli are used in the training as well as shift tasks (Buss, 1953; Eimas, 1965; Slamecka, 1968; Wolff, 1967). This brings into play the potential confounding effects of partial reinforcement. There is 0% reinforcement of responses to the previously correct stimuli in the R shift, but approximately 50% reinforcement of responses to the previously correct stimuli in the NR shift. This fortuitous intermittent reinforcement during the NR shift may have rendered the relearning of the NR shift more difficult than the R shift where there is consistent nonreinforcement of previously correct responses. Thus, this state of affairs may cloud the effects of the theoretical variables affecting the relative R/NR speeds.

In defence of this methodological criticism, two points can be considered. The first is that the assumption of partial reinforcement during NR holds only for subjects who have learned the two stimulus pairs in the original discrimination as one problem, in which case it is true that partial reinforcement of the acquired dimensional responding occurs with only one pair of stimuli during NR. However, if the subject learns the original stimulus pairs as two separate problems, then the situation during NR would not be partial reinforcement as such, but one of 100% reinforcement for one problem and 0% reinforcement for the other. Thus, the basic hypothesis regarding the relative R/NR speeds as a function of dimensional control should not be vitally affected by this criticism. The second point is that inspite of the potential weakness related to the paradigm itself which is constant for all treatment groups, one still has to account for the differential effects due to the treatments. It is justified, however, to view interpretations and conclusions drawn from this study as tentative and as providing leads for further investigation.

In the light of this discussion, the following suggestions are offered for future research. One could make use of what has been described as the "total change design" (Slamecka, 1968). This involves the comparisons between an intradimensional and an extradimensional shift. In this design, totally new cue values appear on the shift problem while the same dimensions as in the original discrimination are retained. For instance, referring to Figure 1 again, instead of the original colors red and green, and the original forms triangle and square, other values on the color dimension, such as yellow and blue, and on the form dimension, such as circle and cross may be used for the shift phase. This design would eliminate the potential interferences due to partial reinforcement, while at the same time would allow the assessment of similar hypotheses as proposed in this study.

Because the present results related errorless learning to dimensional control, a more direct measure of dimensional responding may be called for. This possibility may be afforded by the insertion of "probe" trials during the course of errorless learning. The probe trials would consist of nonfaded stimuli with the same relevant dimension as the main-task stimuli but with different irrelevant dimensions. Consistent responding to the member of the probe stimuli which are on the same relevant dimension as the original stimuli would be evidence for dimensional control.

One final point, though not central to the present thesis, may be noted. This refers to the performance of control subjects in Experiment I. It was found that the NF preschoolers did not complete the NR shift faster than the R shift. In fact, the evidence was for a faster R than NR learning. This particular finding is inconsistent with developmental hypotheses which specify a relative R/NR speed ontogeny. As noted earlier, Kendler and Kendler (1968, 1970), and Tighe and Tighe (1969, 1970) are proponents of such a theory. While it was not the purpose of this study to test this aspect of these theories, it is interesting to note that the present finding is one of a growing number of studies, independent of those done by the Kendlers and the Tighes themselves, which report either no R/NR differential or actually a superior R in preschool children (e.g. Caron, 1969, 1970; Cole, 1973; Dickerson, 1966; Wolff, 1967). This problem is an issue by itself and is beyond the scope of the present study. Nevertheless, since the procedure employed for the control subjects in this study is directly comparable to those used in discrimination shift experiments described in the literature, it may be regarded as another case where the predicted age-related phenomenon did not occur. One should further note that the failure to obtain NR superiority with preschool children does not necessarily involve a fundamental disagreement with the basic tenets of these theories which may simply have underestimated the capacity of younger children to learn on a dimensional basis. As far as this study is concerned, the objective of including conventionally trained subjects was to provide a control group to which the errorless groups could be compared. It appears that this purpose has been served in this study as interpretable differences were found between the control and the errorless groups.

References

Acker, L. E. Errorless discrimination training in autistic and normal children. Unpublished doctoral dissertation, University of California, Los Angeles, 1966.

Bricker, W. A., Heal, L. W., Bricker, D. D., Hayes, W. A., & Larsen, L. A. Discrimination learning and learning set with institutionalized retarded children. American Journal of Mental Deficiency, 1969, 74, 242-248.

Buss, A. H. Rigidity as a function of reversal and nonreversal shifts in the learning of successive discriminations. Journal of Experimental Psychology, 1953, 45, 75-81.

Buss, A. H. Reversal and nonreversal shifts in concept formation with partial reinforcement eliminated. Journal of Experimental Psychology, 1956, 52, 162-166.

Campione, J. C. Optional intradimensional and extradimensional shifts in children as a function of age. Journal of Experimental Psychology, 1970, 84, 296-300.

Campione, J. C. The performance of preschool children on reversal and two types of extradimensional shifts. Journal of Experimental Child Psychology, 1971, 11, 480-490.

Caron, A. J. Discrimination shifts in three-year-olds as a function of dimensional salience. Developmental Psychology, 1969, 1, 333-339.

Cheney, T., & Stein, N. Fading procedures and oddity learning in kindergarten children. Journal of Experimental Child Psychology, 1974, 17, 313-321.

Cole, M. A developmental study of factors influencing discrimination transfer. Journal of Experimental Child Psychology, 1973, 16, 126-147.

Cole, R. E., Dent, H. E., Eguchi, P. E., Fujii, K. K., & Johnson, R. C. Transposition with minimal errors during training trials. Journal of Experimental Child Psychology, 1964, 1, 355-359.

Dehn, J. K. An investigation of the development and maintenance of the negative behavior of autistic children. Unpublished doctoral dissertation, Washington University, St. Louis, 1970.

Dickerson, D. J. Performance of preschool children on three discrimination shifts. Psychonomic Science, 1966, 4, 417-418.

Dickerson, D. J. Irrelevant stimulus dimensions and dimensional transfer in the discrimination learning of children. Journal of Experimental Child Psychology, 1967, 5, 228-236.

Dickerson, D. J., Novik, N., & Gould, S. A. Acquisition and extinction rates as determinants of age changes in discrimination shift behavior. Journal of Experimental Psychology, 1972, 95, 116-122.

Dickerson, D. J., Wagner, J. F., & Campione, J. C. Discrimination shift performance of kindergarten children as a function of variation of the irrelevant shift dimension. Developmental Psychology, 1970, 3, 229-235.

Ehrenfreund, D. An experimental test of the continuity theory of discrimination learning with pattern vision. Journal of Comparative and Physiological Psychology, 1948, 41, 408-422.

Eimas, P. D. Comment: comparisons of reversal and nonreversal shifts. Psychonomic Science, 1965, 3, 445-446.

Eimas, P. D. Observing response, attention, and the overtraining reversal effect. Journal of Experimental Psychology, 1969, 82, 499-502.

Fellows, B. J. Chance stimulus sequences for discrimination tasks. Psychological Bulletin, 1967, 67, 87-92.

Fisher, M. A., & Zeaman, D. An attention-retention theory of retardate discrimination learning. In N. R. Ellis (Ed.), International Review of Research in Mental Retardation (Vol. 6). New York: Academic Press, 1973.

Gibson, J. J. The perception of the visual world. Boston: Houghton Mifflin, 1950.

Gibson, J. J. Perception as a function of stimulation. In S. Koch (Ed.), Psychology: A study of a science (Vol. 1). New York: McGraw-Hill, 1959.

Gollin, E. S., & Savoy, P. Fading procedures and conditional discrimination in children. Journal of the Experimental Analysis of Behavior, 1968, 11, 443-451.

Hanson, H. M. Effects of discrimination training on stimulus generalization. Journal of Experimental Psychology, 1959, 58, 321-334.

Harlow, H. F. Analysis of discrimination learning by monkeys. Journal of Experimental Psychology, 1950, 40, 26-39.

Harlow, H. F. Learning set and error factor theory. In S. Koch (Ed.), Psychology: A study of a science (Vol. 2). New York: McGraw-Hill, 1959.

Harrow, M., & Friedman, G. B. Comparing reversal and nonreversal shifts in concept formation with partial reinforcement controlled. Journal of Experimental Psychology, 1958, 55, 592-598.

Hewett, F. M., Mayhew, D., & Rabb, E. An experimental reading program for neurologically impaired, mentally retarded, and severely emotionally disturbed children. American Journal of Orthopsychiatry, 1967, 37, 35-48.

House, B., & Zeaman, D. Reversal and nonreversal shifts in discrimination learning in retardates. Journal of Experimental Psychology, 1962, 63, 444-451.

Hull, C. L. Simple trial-and-error learning: a study in psychological theory. Psychological Review, 1930, 37, 241-256.

Hull, C. L. Principles of behavior. New York: Appleton-Century-Crofts, 1943.

James, W. Principles of psychology. New York: Holt, 1890.

Jenkins, H. M. The effects of discrimination training on extinction. Journal of Experimental Psychology, 1961, 61, 111-121.

Jenkins, H. M. Generalization gradients and the concept of inhibition. In D. I. Mostofsky (Ed.), Stimulus generalization. Stanford: Stanford University Press, 1965.

Karraker, R. J., & Doke, L. A. Errorless discrimination of alphabet letters: effects of time and method of introducing competing stimuli. Journal of Experimental Education, 1970, 38, 29-35.

Keller, F. S., & Schoenfeld, W. N. Principles of psychology. New York: Appleton-Century-Crofts, 1950.

Kendler, H. H., & D'Amato, M. F. A comparison of reversal shifts and nonreversal shifts in human concept formation behavior. *Journal of Experimental Psychology*, 1955, *49*, 165-174.

Kendler, H. H., & Kendler, T. S. Vertical and horizontal processes in problem solving. *Psychological Review*, 1962, *69*, 1-16.

Kendler, H. H., & Kendler, T. S. Mediation and conceptual behavior. In J. T. Spence (Ed.), *Psychology of learning and motivation* (Vol. 2). New York: Academic Press, 1968.

Kendler, H. H., & Kendler, T. S. Developmental processes in discrimination learning. *Human Development*, 1970, *13*, 65-89.

Kendler, T. S. Development of mediating responses in children. *Monographs of the Society for Research in Child Development*, 1963, *28*, 33-52.

Kirk, R. E. *Experimental design: Procedures for the behavioral sciences*. Belmont: Wadsworth, 1968.

Krechevsky, I. A study of the continuity of the problem-solving process. *Psychological Review*, 1938, *45*, 107-133.

Lawrence, D. H. Acquired distinctiveness of cues:II. Selective association in a constant stimulus situation. *Journal of Experimental Psychology*, 1950, *40*, 175-188.

Lawrence, D. H. The transfer of a discrimination along a continuum. *Journal of Comparative and Physiological Psychology*, 1952, *45*, 511-516.

Lee, L. C. Concept utilization in preschool children. *Child Development*, 1965, *36*, 221-227.

Moore, R., & Goldiamond, I. Errorless establishment of visual discrimination using fading procedures. *Journal of the Experimental Analysis of Behavior*, 1964, *7*, 269-272.

Pavlov, I. P. *Conditioned reflexes*. London: Oxford University Press, 1927.

Pierrel, R., & Sherman, J. G. Generalization and discrimination as a function of the S^D-S^Δ intensity difference. *Journal of the Experimental Analysis of Behavior*, 1962, *5*, 67-71.

Powers, R. B., Cheney, C. C., & Agostino, N. R. Errorless training of a visual discrimination in preschool children. The Psychological Record, 1970, 20, 45-50.

Reynolds, G. S. Behavioral contrast. Journal of the Experimental Analysis of Behavior, 1961, 4, 57-71.

Schlosberg, H., & Solomon, R. L. Latency of response in a choice of discrimination. Journal of Experimental Psychology, 1943, 33, 22-39.

Schwartz, S. H., Firestone, I. J., & Terry, S. Fading techniques and concept learning in children. Psychonomic Science, 1971, 25, 83-84.

Shepp, B. E., & Turrisi, F. D. Effects of overtraining on the acquisition of intradimensional and extradimensional shifts. Journal of Experimental Psychology, 1969, 82, 46-51.

Sidman, M., & Stoddard, L. T. The effectiveness of fading in programming a simultaneous form discrimination for retarded children. Journal of the Experimental Analysis of Behavior, 1967, 10, 8-15.

Skinner, B. F. The behavior of organisms: An experimental analysis. New York: Appleton-Century-Crofts, 1938.

Slamecka, N. J. A methodological analysis of shift paradigms in human discrimination learning. Psychological Bulletin, 1968, 69, 423-438.

Solokov, E. N. Neuronal models and the orienting reflex. In M. A. B. Brazier (Ed.), The central nervous system and behavior. New York: Josiah Macy, Jr. Foundation, 1960.

Spence, K. W. The nature of discrimination learning in animals. Psychological Review, 1936, 43, 427-449.

Spence, K. W. The nature of response in discrimination learning. Psychological Review, 1952, 59, 89-93.

Spence, K. W. Behavior theory and conditioning. New Haven: Yale University Press, 1956.

Suchman, R. G., & Trabasso, T. Color and form preference in young children. Journal of Experimental Child Psychology, 1966a, 3, 177-187.

Suchman, R. G., & Trabasso, T. Stimulus preference and cue function in young children's concept attainment. Journal of Experimental Child Psychology, 1966b, 3, 188-198.

Sutherland, N. S. Stimulus analyzing mechanisms. In _Proceedings of a symposium on the mechanization of thought processes_ (Vol. 2). London: Her Majesty's Stationery Office, 1959.

Terrace, H. S. Discrimination learning with and without "errors". _Journal of the Experimental Analysis of Behavior_, 1963a, _6_, 1-27.

Terrace, H. S. Errorless transfer of a discrimination across two continua. _Journal of the Experimental Analysis of Behavior_, 1963b, _6_, 223-232.

Terrace, H. S. Wavelength generalization after discrimination learning with and without errors. _Science_, 1964, _144_, 78-80.

Terrace, H. S. Stimulus control. In W. K. Honig (Ed.), _Operant behavior: Areas of research and application_. New York: Appleton-Century-Crofts, 1966a.

Terrace, H. S. Discrimination learning and inhibition. _Science_, 1966b, _154_, 1677-1680.

Terrace, H. S. Escape from S-. _Learning and Motivation_, 1971, _2_, 148-163.

Tighe, L. S. Effects of perceptual pretraining on reversal and nonreversal shifts. _Journal of Experimental Psychology_, 1965, _4_, 379-385.

Tighe, L. S., & Tighe, T. J. Overtraining and discrimination shift behavior in children. _Psychonomic Science_, 1965, _2_, 365-366.

Tighe, L. S., & Tighe, T. J. Discrimination learning: two views in historical perspective. _Psychological Bulletin_, 1966, _66_, 353-370.

Tighe, T. J., Glick, J., & Cole, M. Subproblem analysis of discrimination shift learning. _Psychonomic Science_, 1971, _24_, 159-160.

Tighe, T. J., & Tighe, L. S. Overtraining and optional shift behavior in rats and children. _Journal of Comparative and Physiological Psychology_, 1966, _62_, 49-54.

Tighe, T. J., & Tighe, L. S. Differentiation theory and concept-shift behavior. _Psychological Bulletin_, 1968, _70_, 756-761.

Tighe, T. J., & Tighe, L. S. Facilitation of transposition and reversal learning in children by prior perceptual training. Journal of Experimental Child Psychology, 1969, 8, 366-374.

Tighe, T. J., & Tighe, L. S. Optional shift behavior of children as a function of age, type of pretraining, and stimulus salience. Journal of Experimental Child Psychology, 1970, 9, 272-285.

Tighe, T. J., & Tighe, L. S. Stimulus control in children's learning. In A. D. Pick (Ed.), Minnesota symposium on child psychology (Vol. 6). Minneapolis: University of Minnesota Press, 1972.

Wolff, J. L. Concept-shift and discrimination-reversal learning in humans. Psychological Bulletin, 1967, 67, 369-408.

Wyckoff, L. R. The role of observing responses in discrimination learning. Psychological Review, 1952, 59, 431-442.

Zeaman, D., & House, B. J. The role of attention in retardate discrimination learning. In N. R. Ellis (Ed.), Handbook of mental deficiency. New York: McGraw-Hill, 1963.

Appendix 1

Tables

Table 1

Summary of the Analysis of Variance on Preshift Trials to Criterion for Preschool Subjects, Transformed Data.

Source	SS	df	MS	F
Training (T)	37.1859	1	37.1859	6.4459*
Shifts (S)	2.3123	1	2.3123	0.4008
Dimension (D)	67.7758	1	67.7758	11.7485**
Sex (X)	0.0002	1	0.0002	0.0000
T x S	2.8375	1	2.8375	0.4919
T x D	59.9190	1	59.9190	10.3866**
S x D	1.3861	1	1.3861	0.2403
T x X	0.3215	1	0.3215	0.0557
S x X	1.5447	1	1.5447	0.2678
D x X	0.0035	1	0.0035	0.0006
T x S x D	3.2765	1	3.2765	0.5680
T x S x X	5.1942	1	5.1942	0.9004
T x D x X	0.4362	1	0.4362	0.0756
S x D x X	0.0218	1	0.0218	0.0038
T x S x D x X	0.1526	1	0.1526	0.0265
Subj. w/cell	276.9060	48	5.7689	-
Total	459.2738	63		

* $p < .05$

** $p < .01$

Table 2

Summary of the Analysis of Variance on Shift Trials to Criterion for Preschool Subjects, Transformed Data.

Source	SS	df	MS	F
Training (T)	52.3344	1	52.3344	5.3858*
Shifts (S)	543.7771	1	543.7771	55.9607**
Dimensions (D)	4.1658	1	4.1658	0.4287
Sex (X)	13.7097	1	13.7097	1.4109
T x S	42.2163	1	42.2163	4.3445*
T x D	15.8047	1	15.8047	1.6265
S x D	1.7620	1	1.7620	0.1813
T x X	0.5673	1	0.5673	0.0584
S x X	1.6860	1	1.6860	0.1735
D x X	0.0129	1	0.0129	0.0013
T x S x D	0.2932	1	0.2932	0.0302
T x S x X	62.2779	1	62.2779	6.4091*
T x D x X	3.2249	1	3.2249	0.3319
S x D x X	1.8206	1	1.8206	0.1874
T x S x D x X	4.2975	1	4.2975	0.4423
Subj. w/cell	466.4216	48	9.7171	-
Total	466.4216	63		

* $p < .05$

** $p < .01$

Table 3

Mean Trials (raw scores) to Shift Criterion of Preschoolers as a Function of Preshift Training and Shift Conditions.

		Shifts		
		Reversal	Nonreversal	Combined
Preshift Training	Fading	22.81	74.63	48.72
	Nonfading	22.13	49.38	35.75
Combined		22.47	62.00	

Table 4

Mean Trials (raw scores) to Shift Criterion of Preschoolers as a Function of Sex, Preshift Training and Shift Conditions

		Males		Females	
		R	NR	R	NR
Preshift Training	Fading	25.75	60.75	19.86	88.50
	Nonfading	16.25	53.63	28.00	45.13

Table 5

Summary of the Analysis of Variance on Shift Trials to Criterion of Preschool Girls, Transformed Data.

Source	SS	df	MS	F
Training (T)	31.8962	1	31.8962	4.4912*
Shift (S)	302.9121	1	302.9121	42.6520**
T x S	103.5217	1	103.5217	14.5765
Subj. w/cell	198.8545	28	7.1020	-
Total	637.1845	31		

* $\underline{p} < .05$
** $\underline{p} < .01$

Table 6

Summary of the Analysis of Variance on Shiht Trials to Criterion for Preschool Boys, Transformed Data.

Source	SS	df	MS	F
Training (T)	21.0049	1	21.0049	1.9673
Shift (S)	242.5505	1	242.5505	22.7172**
T x S	0.9736	1	0.9736	0.0912
Subj. w/cell	298.9541	28	10.6769	-
Total	563.4831	31		

** $\underline{p} < .01$

Table 7

Summary of the Analysis of Variance on Shift Errors to Criterion for Preschool Subjects, Transformed Data.

Source	SS	df	MS	F
Training (T)	48.1289	1	48.1289	11.4748**
Shifts (S)	197.8223	1	197.8223	47.1642**
Dimensions (D)	0.0196	1	0.0196	0.0047
Sex (X)	8.0514	1	8.0514	1.9196
T x S	17.0791	1	17.0791	4.0720*
T x D	9.1658	1	9.1658	2.1853
S x D	0.0151	1	0.0151	0.0036
T x X	2.5440	1	2.5440	0.6065
S x X	0.0514	1	0.0514	0.0123
D x X	1.2377	1	1.2377	0.2951
T x S x D	0.1985	1	0.1985	0.0473
T x S x X	17.9333	1	17.9333	4.2756*
T x D x X	1.1448	1	1.1448	0.2729
S x D x X	0.0352	1	0.0352	0.0084
T x S x D x X	1.2450	1	1.2450	0.2968
Subj. w/cell	201.3277	48	4.1943	-
Total	505.9998	63		

* $p < .05$
** $p < .01$

Table 8

Mean Errors (raw scores) to Shift Criterion of Preschoolers as a Function of Preshift Training and Shift Conditions.

		Shifts		
		Reversal	Nonreversal	Combined
Preshift Training	Fading	8.25	26.81	17.53
	Nonfading	5.94	14.56	10.25
Combined		7.09	20.69	

Table 9

Mean Errors (raw scores) to Shift Criterion of Preschoolers as a Function of Sex, Preshift Training and Shift Conditions.

		Males		Females	
		R	NR	R	NR
Preshift Training	Fading	7.75	21.38	8.75	32.25
	Nonfading	4.25	15.75	7.63	13.38

Table 10

Summary of the Analysis ofof riance on the Proportions Errors during Shift for Preschool Subjects.

Source	SS	df	MS	F
Training (T)	842.2747	1	842.2747	6.8830*
Shift (S)	64.0234	1	64.0234	0.5232
Dimensions (D)	151.0435	1	151.0435	1.2343
Sex (X)	29.7416	1	29.7416	0.2430
T x S	9.1045	1	9.1045	0.0744
T x D	92.1111	1	92.1111	0.7527
S x D	81.0581	1	81.0581	0.6624
T x X	91.1583	1	91.1583	0.7449
S x X	126.2923	1	126.2923	1.0320
D x X	91.7465	1	91.7465	0.7497
T x S x D	2.0796	1	2.0796	0.0170
T x S x X	1.1802	1	1.1802	0.0096
T x D x X	39.8095	1	39.8095	0.3253
S x D x X	108.9271	1	108.9271	0.8901
T x S x D x X	0.0237	1	0.0237	0.0002
Subj. w/cell	5873.8090	48	122.3710	-
Total	7604.3831	63		

* $p < .05$

Table 11

Summary of the Analysis of Variance on the Number of Perseverative Errors on the First Changed Shift Pair, Preschoolers, Transformed Data.

Source	SS	df	MS	F
Training (T)	57.4658	1	57.4658	23.7842**
Dimension (D)	13.3051	1	13.3051	5.5068*
Sex (X)	0.4919	1	0.4919	0.2036
T x D	2.4044	1	2.4044	0.9952
T x X	0.9474	1	0.9474	0.3912
D x X	1.8540	1	1.8540	0.7674
T x D x X	0.1853	1	0.1853	0.0767
Subj. w/cell	135.3033	56	2.4161	-
Total	211.9572	63		

* $p < .05$
** $p < .01$

Table 12

Summary of the Analysis of Variance on Shift Trials to Criterion, Excluding Perseverative Errors on the First Changed Shift Pair, Preschoolers, Transformed Data.

Source	SS	df	MS	F
Training (T)	17.3892	1	17.3892	1.7180
Shift (S)	588.7788	1	588.7788	58.1685**
Dimension (D)	12.7090	1	12.7090	1.2556
Sex (X)	12.3197	1	12.3197	1.2171
T x S	44.0667	1	44.0667	4.3536*
T x D	12.2322	1	12.2322	1.2085
S x D	2.9609	1	2.9609	0.2925
T x X	0.0264	1	0.0264	0.0026
S x X	4.3692	1	4.3692	0.4317
D x X	0.0950	1	0.0950	0.0094
T x S x D	0.0729	1	0.0729	0.0072
T x S x X	78.6655	1	78.6655	7.7718**
T x D x X	3.9998	1	3.9998	0.3952
S x D x X	3.4479	1	3.4479	0.3406
T x S x D x X	5.0187	1	5.0187	0.4958
Subj. w/cell	485.8538	48	10.1220	-
Total	1272.0057	63		

* $p < .05$

** $p < .01$

Table 13

Summary of the Analysis of Variance on Shift Trials to Criterion, Excluding Perseverative Errors on the First Changed Shift Pair, Preschool Girls, Transformed Data.

Source	SS	df	MS	F
Training (T)	9.3571	1	9.3571	1.4066
Shift (S)	347.2798	1	347.2798	52.2035**
T x S	120.3420	1	120.3420	18.0900**
Subj. w/cell	186.2678	28	6.6524	-
Total	663.2467	31		

** p < .01

Table 14

Summary of the Analysis of Variance on Shift Trials to Criterion, Excluding Perseverative Errors on the First Changed Shift Pair, Preschool Boys, Transformed Data.

Source	SS	df	MS	F
Training (T)	8.0361	1	8.0361	0.6614
Shift (S)	246.1096	1	246.1096	20.2566**
T x S	2.5008	1	2.5008	0.2058
Subj. w/cell	340.1895	28	12.1496	-
Total	596.8360	31		

** p < .01

Table 15

Mean Trials (raw scores) to Shift Criterion of Preschoolers as a Function of Sex, Preshift Training and Shift Conditions, Excluding Perseverative Errors on the First Changed Shift Pair.

		Males		Females	
		R	NR	R	NR
Preshift Training	Fading	21.50	52.75	13.00	80.25
	Nonfading	14.13	52.13	26.25	43.38

Table 16

Number of Preschool Spontaneous Shifters as a Function of Preshift Training Methods.

		Preshift Training		
		Fading	Nonfading	Total
Spontaneous Shifting	yes	17	18	35
	No	15	15	29
Total		32	32	64

Table 17

Number of Preschool Spontaneous Shifters as a Function of Original Relevant Dimensions.

		Original Dimension		
		Form	Color	Total
Spontaneous Shifting	Yes	20	15	35
	No	12	17	29
Total		32	32	64

Table 18

Number of Preschool Spontaneous Shifters as a Function of Preshift Training Methods and Relevant Dimensions.

		Fading		Nonfading		
		Form	Color	Form	Color	Total
Spontaneous	Yes	9	8	11	7	35
Shifting	No	7	8	5	9	29
Total		16	16	16	16	64

Table 19

Summary of the Analysis of Variance on the Proportions Correct of the R versus NR-C Subproblems for Preschool Subjects.

Source	SS	df	MS	F
Training (T)	0.0065	1	0.0065	0.0318
Subproblem (P)	7.5657	1	7.5657	37.0677**
Sex (X)	0.6090	1	0.6090	2.9840
T x P	0.0747	1	0.0747	0.3660
T x X	0.0189	1	0.0189	0.0924
P x X	0.1892	1	0.1892	0.9269
T x P x X	0.3574	1	0.3574	1.7510
Subj. w/cell	11.4299	56	0.2041	-
Total	20.2513	63		

** $p < .01$

Table 20

Summary of the Analysis of Variance on the Proportions Correct of the R versus NR-U Subproblems for Preschool Subjects.

Source	SS	df	MS	F
Training (T)	0.0070	1	0.0070	0.0400
Subproblem (P)	1.9239	1	1.9239	11.0534**
Sex (X)	0.0097	1	0.0097	0.0557
T x P	0.1913	1	0.1913	1.0990
T x X	0.0342	1	0.0342	0.1967
P x X	0.0610	1	0.0610	0.3503
T x P x X	0.3026	1	0.3026	1.7387
Subj. w/cell	9.7470	56	0.1741	-
Total	12.2767	63		

** $p < .01$

Table 21

Summary of the Analysis of Variance on the Proportions Correct of the NR-C versus NR-U Subproblems for Preschool Subjects.

Source	SS	df	MS	F
Training (T)	0.1273	1	0.1273	0.4553
Sex (X)	0.2846	1	0.2846	1.0181
T x X	0.1704	1	0.1704	0.6094
Subj. w/groups	7.8280	28	0.2796	-
Subproblems (P)	1.8593	1	1.8593	5.3200*
T x P	0.0269	1	0.0269	0.0769
X x P	0.4650	1	0.4650	1.3305
T x X x P	0.0023	1	0.0023	0.0066
P x Subj. w/groups	9.7855	28	0.3495	-
Total	20.5493	63		

* $p < .05$

Table 22

Average Subproblem Correct Proportions, Preschool Subjects.

	Fading			Nonfading		
Sex	R	NR-U	NR-C	R	NR-U	NR-C
Male	80.59	63.75	60.00	84.00	65.71	55.60
Female	84.17	61.25	41.25	70.94	66.43	45.89
Combined	82.38	62.50	50.63	77.47	66.07	50.75

Table 23

Summary of the Analysis of Variance on Preshift Trials to Criterion for Second-grade Subjects, Transformed Data.

Source	SS	df	MS	F
Training (T)	2.6245	1	2.6245	0.3867
Shift (S)	2.8139	1	2.8139	0.4146
Dimension (D)	29.7565	1	29.7565	4.3839*
Sex (X)	10.1761	1	10.1761	1.4992
T x S	1.2769	1	1.2769	0.1881
T x D	38.1620	1	38.1620	5.6223*
S x D	24.6024	1	24.6024	3.6246
T x X	8.0229	1	8.0229	1.1820
S x X	0.7834	1	0.7834	0.1154
D x X	7.8827	1	7.8827	1.1613
T x S x D	16.1797	1	16.1797	2.3837
T x S x X	0.0315	1	0.0315	0.0046
T x D x X	4.4727	1	4.4727	0.6589
S x D x X	3.5430	1	3.5430	0.5220
T x S x D x X	5.0406	1	5.0406	0.7426
Subj. w/cell	325.8062	48	6.7876	-
Total	481.1750	63		

* $p < .05$

Table 24

Summary of the Analysis of Variance on Shift Trials to Criterion for Second-grade Subjects, Transformed Data.

Source	SS	df	MS	F
Training (T)	40.1081	1	40.1081	3.7803
Shift (S)	441.1743	1	441.1743	41.5816**
Dimension (D)	9.7649	1	9.7649	0.9204
Sex (X)	6.2203	1	6.2203	0.5863
T x S	67.7476	1	67.7476	6.3853*
T x D	9.1639	1	9.1639	0.8637
S x D	1.3848	1	1.3848	0.1305
T x X	0.1392	1	0.1392	0.0131
S x X	0.0983	1	0.0983	0.0093
D x X	12.9487	1	12.9487	1.2204
T x S x D	2.8957	1	2.8957	0.2729
T x S x X	3.6079	1	3.6079	0.3401
T x D x X	0.3166	1	0.3166	0.0298
S x D x X	6.8784	1	6.8784	0.6483
T x S x D x X	2.6946	1	2.6946	0.2540
Subj. w/cell	509.2725	48	10.6098	-
Total	1114.4158	63		

* $p < .05$
** $p < .01$

Table 25

Mean Trials (raw scores) to Shift Criterion of Second-graders as a Function of Preshift Training and Shift Conditions.

		Shifts		
		Reversal	Nonreversal	Combined
Preshift Training	Fading	19.44	65.63	42.53
	Nonfading	21.75	42.31	32.03
Combined		20.59	53.97	

Table 26

Summary of the Analysis of Variance on Shift Errors to Criterion for Second-graders, Transformed Data.

Source	SS	df	MS	F
Training (T)	22.7886	1	22.7886	3.6142
Shift (S)	190.0604	1	109.0604	30.1426**
Dimension (D)	6.8709	1	6.8709	1.0897
Sex (X)	4.1974	1	4.1974	0.6657
T x S	9.8989	1	9.8989	1.5699
T x D	2.8014	1	2.8014	0.4443
S x D	0.2174	1	0.2174	0.0345
T x X	0.8533	1	0.8533	0.1353
S x X	1.2404	1	1.2404	0.1967
D x X	7.7353	1	7.7353	1.2268
T x S x D	5.1586	1	5.1586	0.8181
T x S x X	0.7077	1	0.7077	0.1122
T x D x X	1.1691	1	1.1691	0.1854
S x D x X	0.9775	1	0.9775	0.1550
T x D x S x X	0.3858	1	0.3858	0.0612
Subj. w/cell	302.6580	48	6.3054	-
Total	557.7207	63		

** $p < .01$

Table 27

Mean Errors (raw scores) to Shift Criterion of Second-graders as a Function of Preshift and Shift Conditions.

		Shifts		
		Reversal	Nonreversal	Combined
Preshift Training	Fading	5.69	19.69	12.69
	Nonfading	4.19	13.56	8.88
Combined		4.94	16.63	

Table 28

Summary of the Analysis of Variance on the Number of Perseverative Errors on the First Changed Shift Pair, Second-graders, Transformed Data.

Source	SS	df	MS	F
TRAINING (T)	10.1610	1	10.1610	5.4288*
Dimension (D)	2.5444	1	2.5444	1.3594
Sex (X)	0.0070	1	0.0070	0.0037
T x D	1.0391	1	1.0391	0.5552
T x X	0.0000	1	0.0000	0.0000
D x X	0.0025	1	0.0025	0.0013
T x D x X	0.8174	1	0.8174	0.4367
Subj. w/cell	104.8141	56	1.8717	-
Total	119.3855	63		

* $p < .05$

Table 29

Summary of the Analysis of Variance on Shift Trials to Criterion, Excluding Perseverative Errors on the First Changed Shift Pair, Second-graders, Tranformed Data.

Source	SS	df	MS	F
Training (T)	23.7775	1	23.7775	2.2679
Shift (S)	478.3418	1	478.3418	45.6248**
Dimension (D)	3.9057	1	3.9057	0.3725
Sex (X)	4.0050	1	4.0050	0.3820
T x S	78.4431	1	78.4431	7.4820**
T x D	16.7590	1	16.7590	1.5985
S x D	3.8367	1	3.8367	0.3660
T x X	0.0004	1	0.0004	0.0000
S x X	1.2448	1	1.2448	0.1187
D x X	11.1639	1	11.1639	1.0648
T x S x D	0.5576	1	0.5576	0.0532
T x S x X	8.6399	1	8.6399	0.8241
T x D x X	0.0085	1	0.0085	0.0008
S x D x X	15.2096	1	15.2096	1.4507
T x S x D x X	6.7003	1	6.7003	0.6391
Subj. w/cell	503.2434	48	10.4842	-
Total	1155.8372	63		

** $p < .01$

Table 30

Number of Second-grade Spontaneous Shifters as a Function of Preshift Training Methods.

		Preshift Training		
		Fading	Nonfading	Total
Spontaneous	Yes	21	25	46
Shifting	No	11	7	18
Total		32	32	64

Table 31

Number of Second-grade Spontaneous Shifters as a Function of Preshift Training Methods and Relevant Dimensions.

		Fading		Nonfading		
		Form	Color	Form	Color	Total
Spontaneous	Yes	10	11	12	13	46
Shifting	No	6	5	4	3	18
Total		16	16	16	16	64

Table 32

Summary of the Analysis of Variance on the Proportions Correct of the R versus NR-C Subproblems for Second-grade Subjects.

Source	SS	df	MS	F
Training (T)	0.1220	1	0.1220	0.7312
Subproblem (P)	8.2887	1	8.2887	49.6651**
Sex (X)	0.0104	1	0.0104	0.0623
T x P	0.1633	1	0.1633	0.9786
T x X	0.0005	1	0.0005	0.0031
P x X	0.0039	1	0.0039	0.0233
T x P x X	0.1812	1	0.1812	1.0858
Subj. w/cell	9.3460	56	0.1669	-
Total	18.1160	63		

** $p < .01$

Table 33

Summary of the Analysis of Variance on the Proportions Correct of the R versus NR-U Subproblems for Second-grade Subjects.

Source	SS	df	MS	F
Training (T)	0.0661	1	0.0661	0.3451
Subproblem (P)	2.1357	1	2.1357	11.1469**
Sex (X)	0.0057	1	0.0057	0.0297
T x P	0.0973	1	0.0973	0.5078
T x X	0.0008	1	0.0008	0.0043
P x X	0.0079	1	0.0079	0.0412
T x P x X	0.1862	1	0.1862	0.9721
Subj. w/cell	10.7292	56	0.1916	-
Total	13.2289	63		

** $p < .01$

Table 34

Summary of the Analysis of Variance on the Proportions Correct of the NR-C versus NR-U Subproblems for Second-grade Subjects.

Source	SS	df	MS	F
Training (T)	0.4373	1	0.4373	1.0837
Sex (X)	0.0002	1	0.0002	0.0004
T x X	0.2064	1	0.2064	0.5115
Subj. w/groups	11.2980	28	0.4035	-
Subproblems (P)	2.0097	1	2.0097	16.7724**
T x P	0.0085	1	0.0085	0.0710
X x P	0.0007	1	0.0007	0.0058
T x X x P	0.0001	1	0.0001	0.0003
P x Subj. w/groups	3.3549	28	0.1198	-
Total	17.3158	63		

** $p < .01$

Table 35

Average Subproblem Correct Proportions, Second-grade Subjects.

	Fading			Nonfading		
Sex	R	NR-U	NR-C	R	NR-U	NR-C
Male	85.71	61.25	43.75	80.63	70.48	56.83
Female	78.98	66.11	48.48	83.59	67.08	52.09
Combined	82.34	63.68	46.12	82.11	68.78	54.46

Appendix 2

Figures

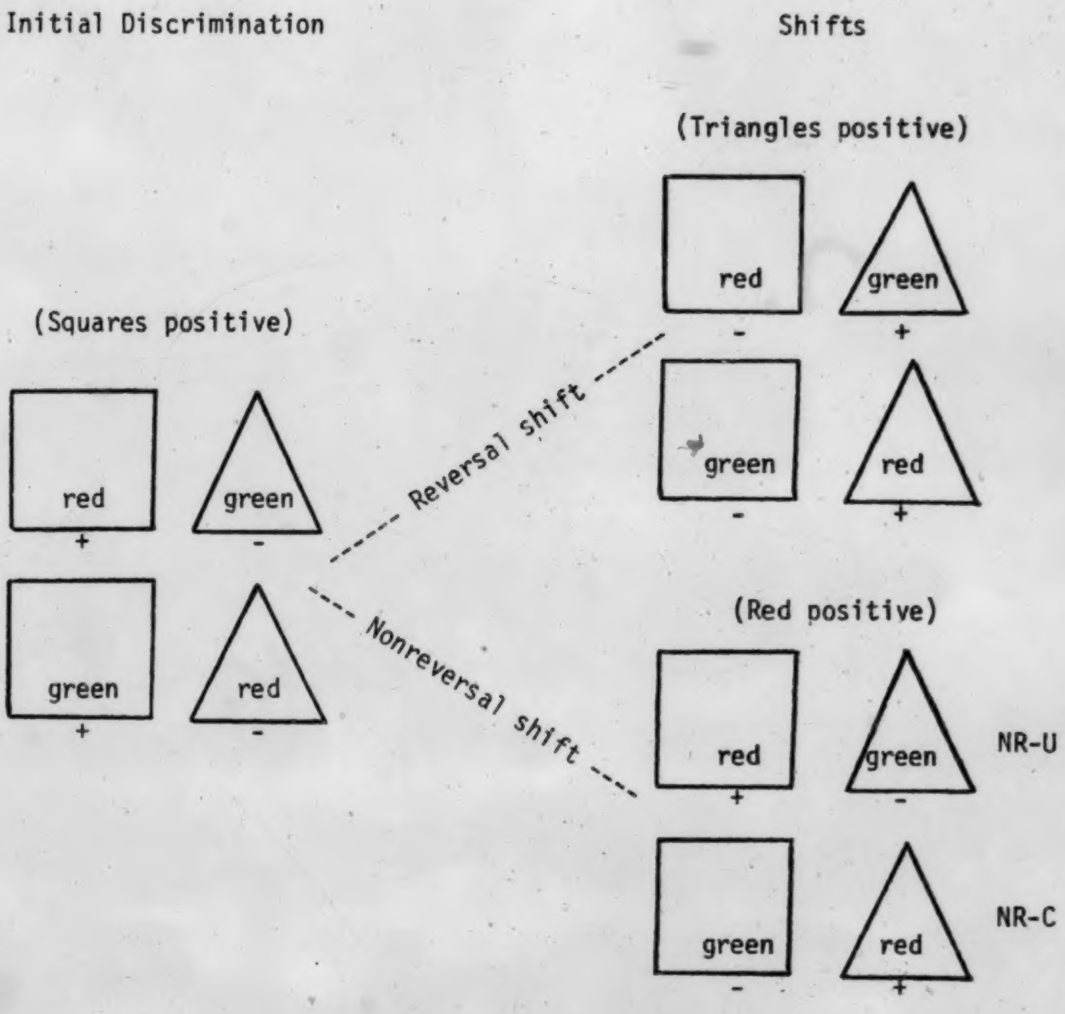

Figure 1. An example of the reversal and nonreversal shifts.
NR-U -- the nonreversal, unchanged pair;
NR-C -- the nonreversal, changed pair.

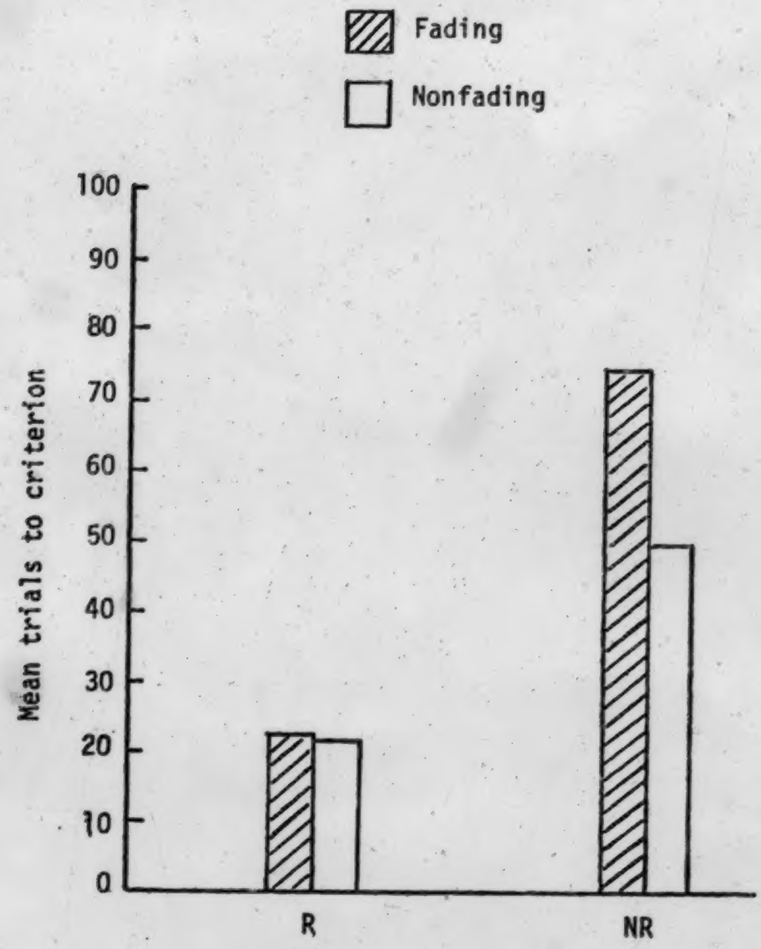

Figure 2. Mean trials (raw scores) to shift criterion of preschool children as a function of Training and Shift conditions.

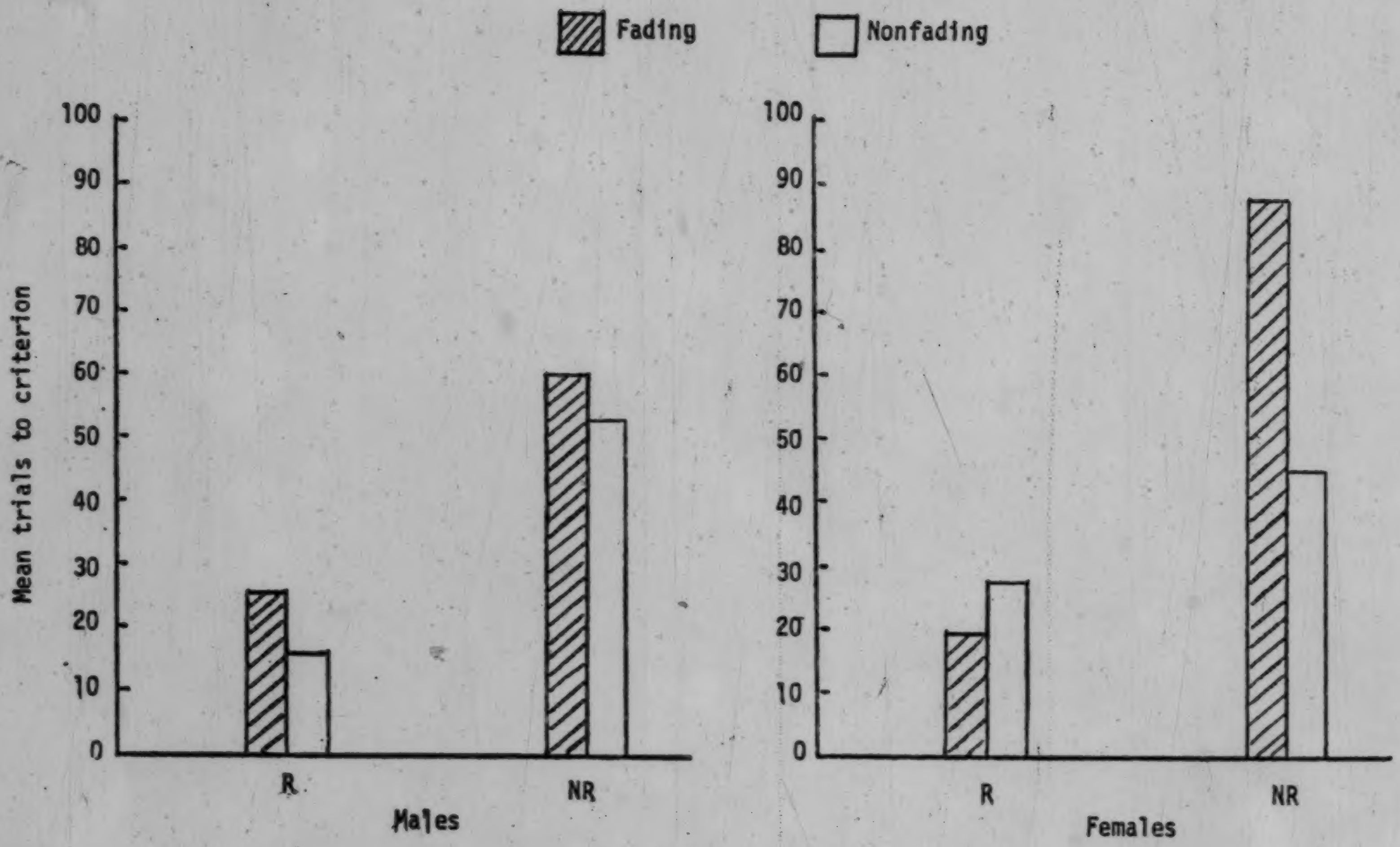

Figure 3. Mean trials (raw scores) to shift criterion of preschool children as a function of Training, Shift, and Sex.

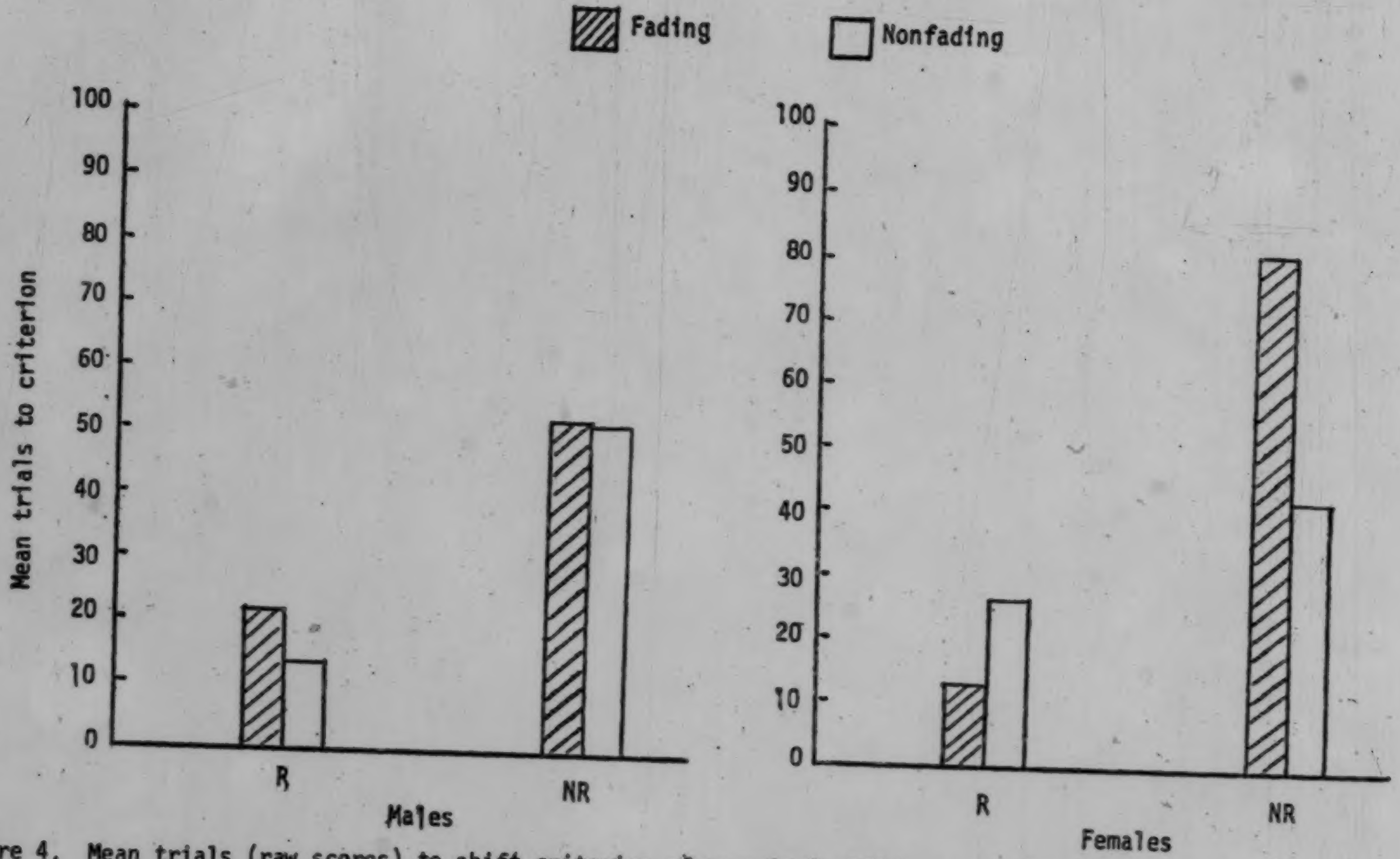

Figure 4. Mean trials (raw scores) to shift criterion of preschool children as a function of Training, Shift, and Sex, excluding multiple errors on the first changed postshift pair.

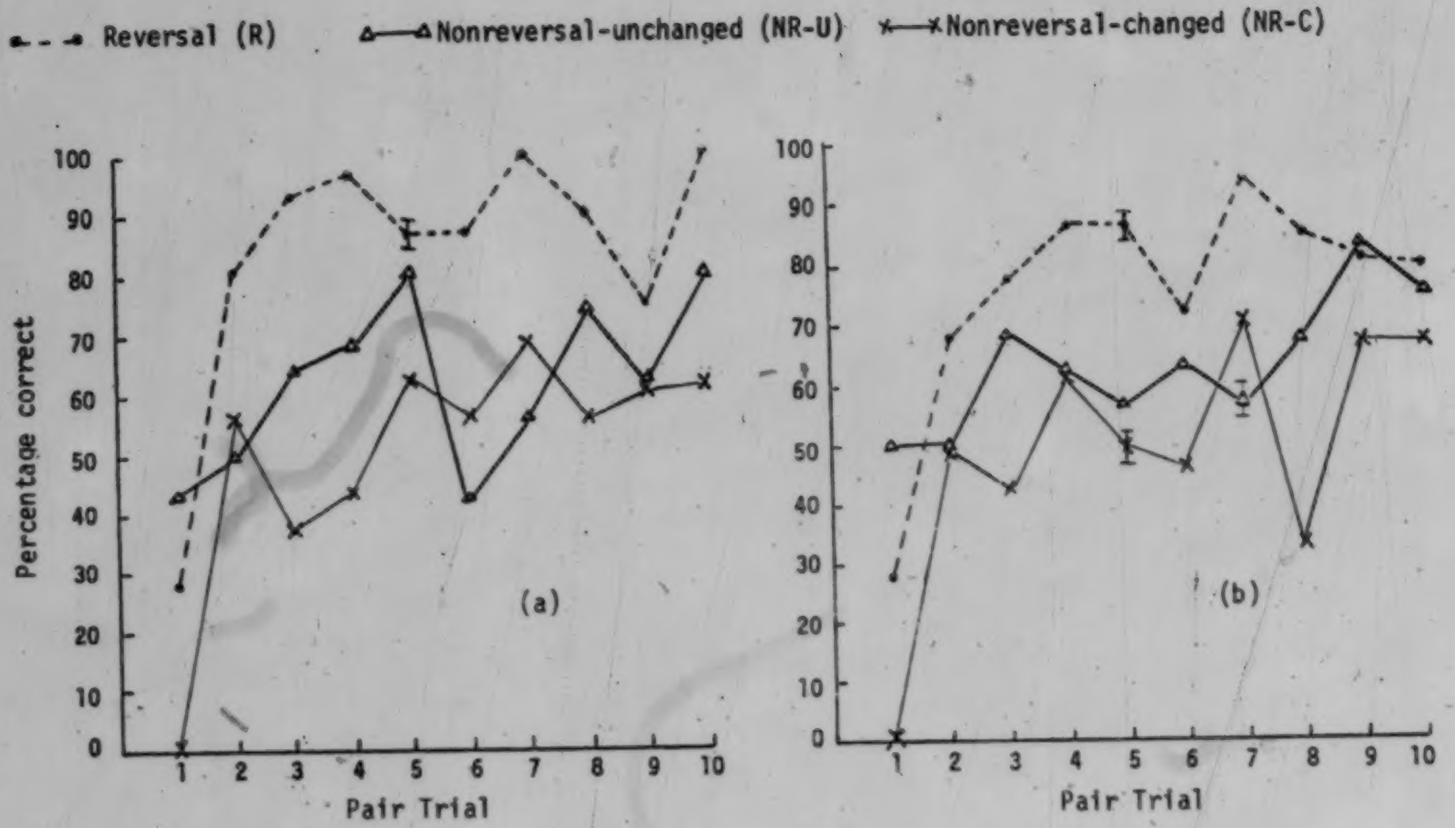

Figure 5. Subproblem learning curves of preschool children. Panel (a) presents data from the fading group and Panel (b) from the nonfading group. I indicates last trial with a full slate of Ss.

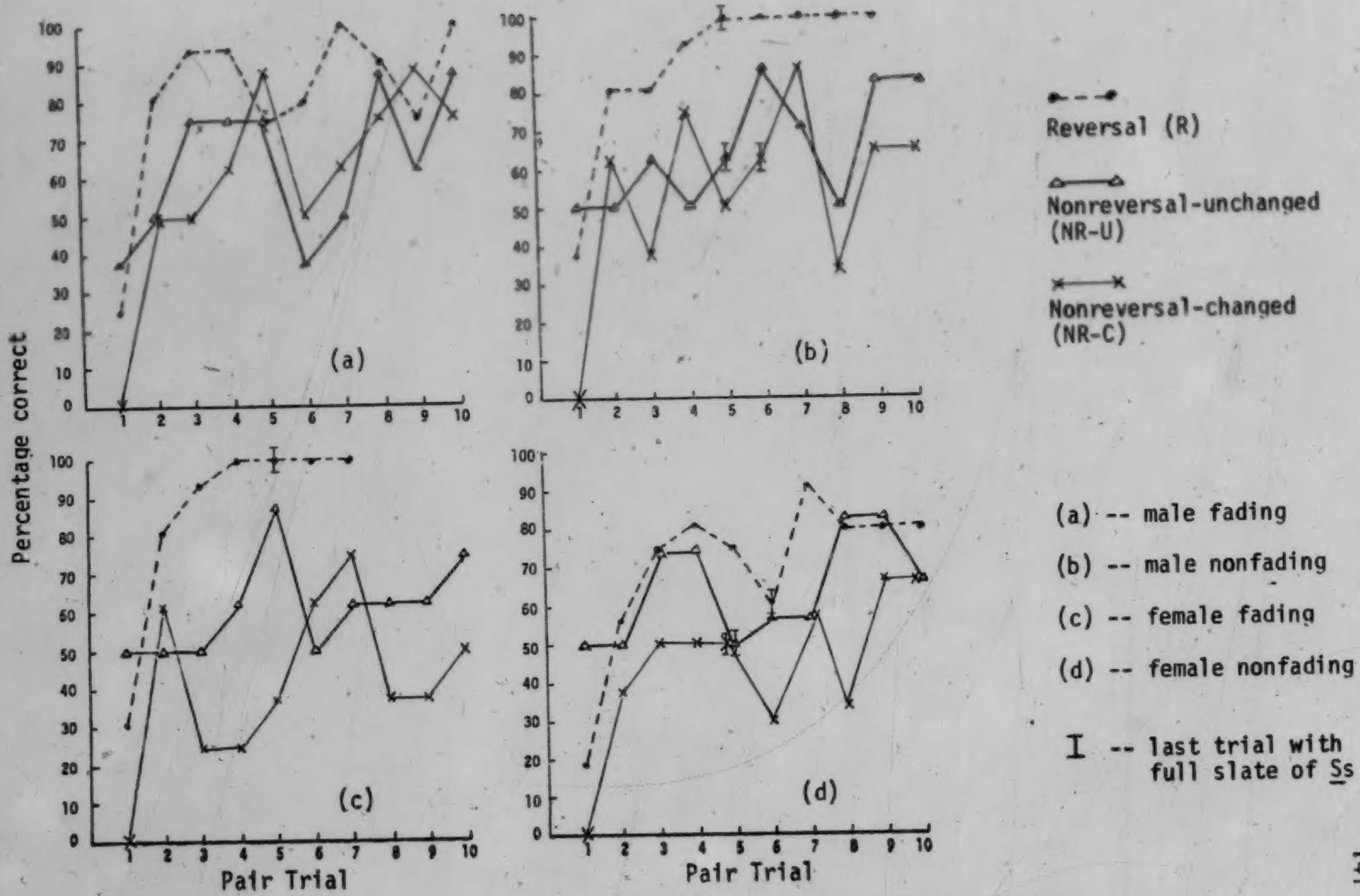

Figure 6. Subproblem learning curves of preschoolers as a function of Training and Sex.

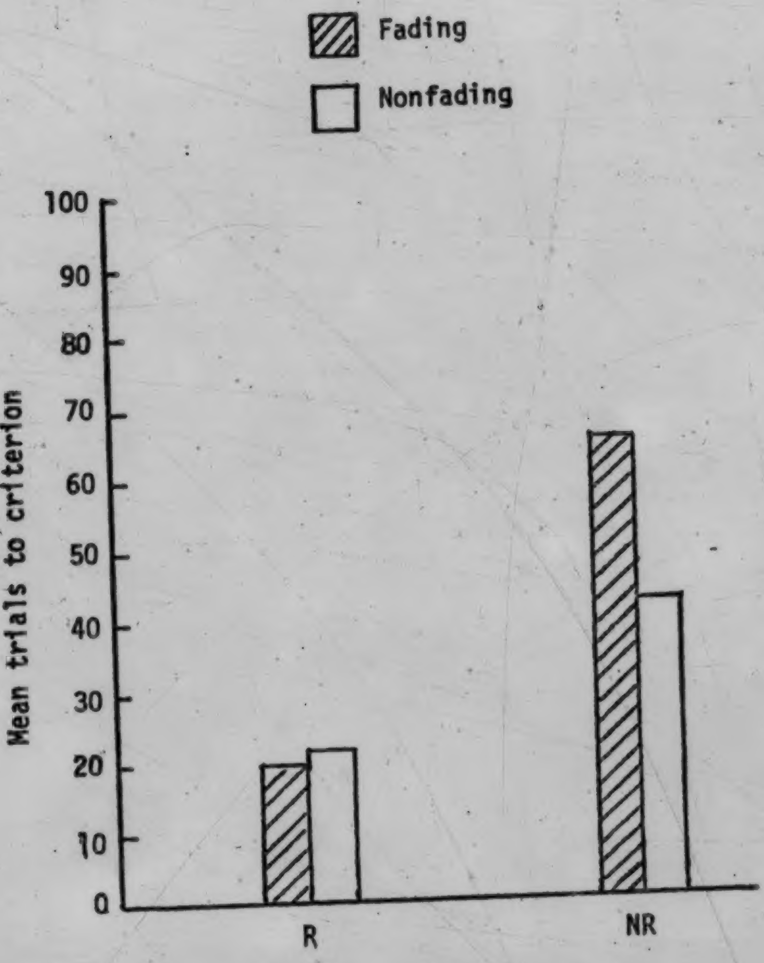

Figure 7. Mean trials (raw scores) to shift criterion of second-grade children as a function of Training and Shift conditions.

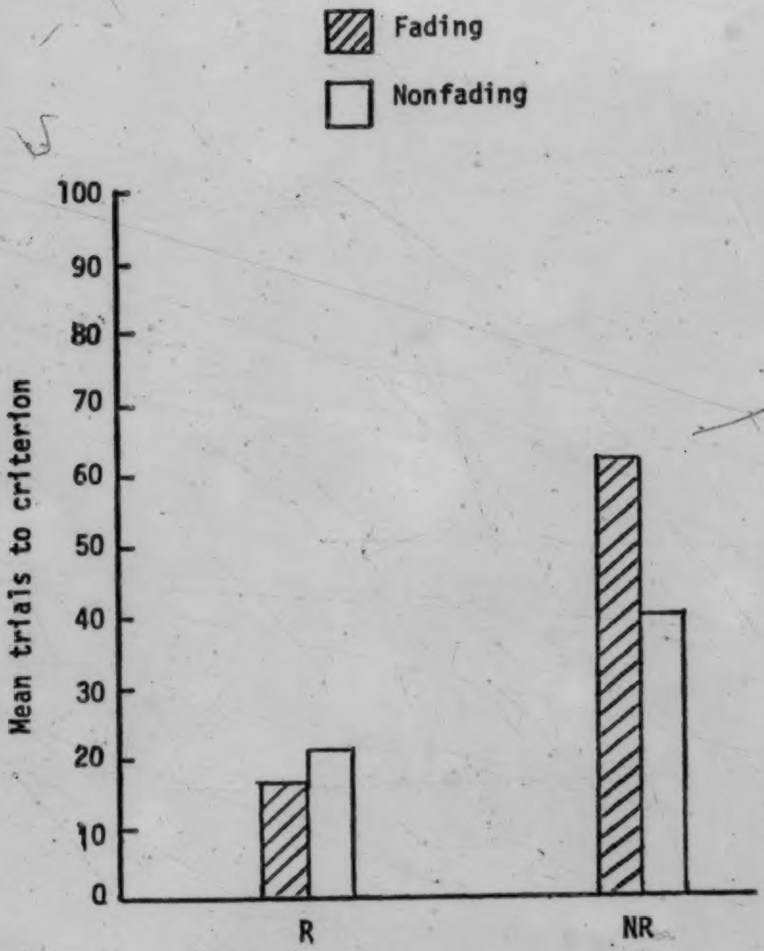

Figure 8. Mean trials (raw scores) to shift criterion of second-grade children as a function of Training and Shift conditions, excluding perseverative errors on the first changed postshift pair.

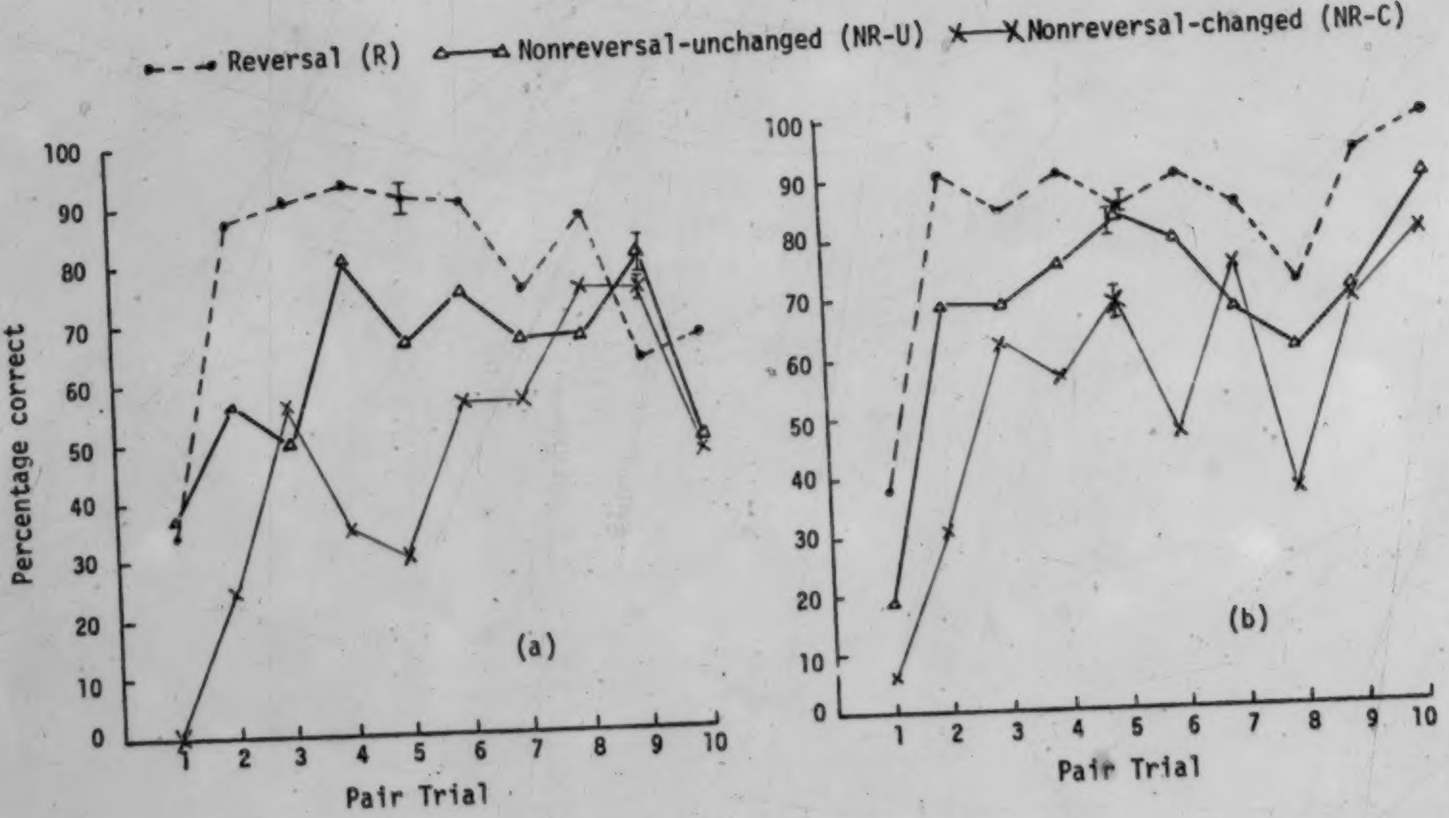

Figure 9. Subproblem learning curves of second-grade children. Panel (a) presents data from the fading group and Panel (b) from the nonfading group. I indicates last trial with a full slate of S̲s̲.

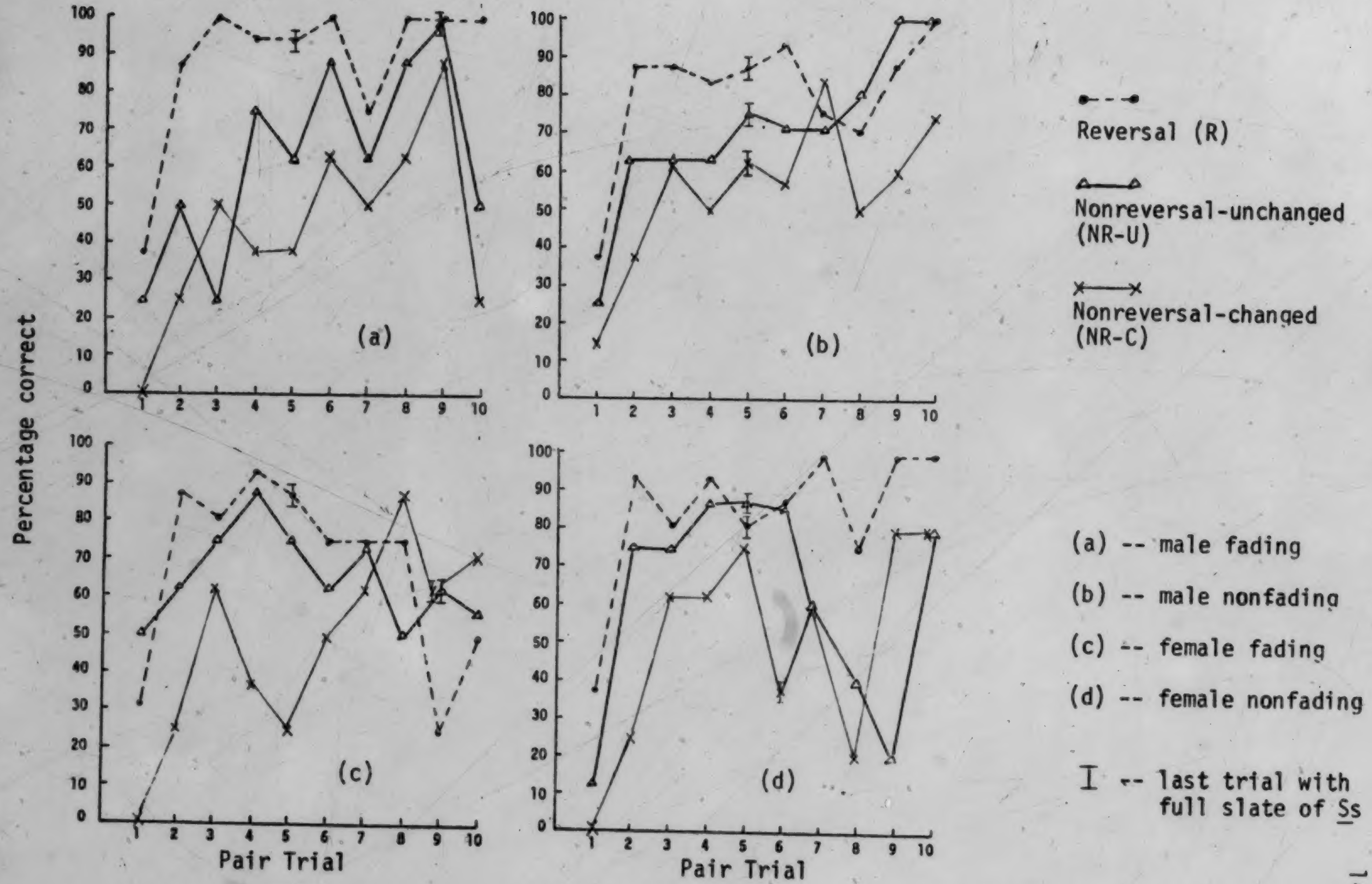

Figure 10. Subproblem learning curves of second-graders as a function of Training and Sex.

Appendix 3

Arrangement of Stimulus Presentations

Initial Discrimination		Shift			
		Reversal		Nonreversal	
Relevant Dimension	Positive Cue	Relevant Dimension	Positive Cue	Relevant Dimension	Positive Cue
Form	Square	Form	Triangle	Color	Red or Green
	Triangle		Square		Red or Green
Color	Red	Color	Green	Form	Square or Triangle
	Green		Red		Square or Triangle

The above shows the relevant dimensions and their respective positive cues appropriate for solution in the initial discrimination and shift tasks. Consider, for example, a subject who had Form as the relevant dimension and Square as the positive cue during initial discrimination. In a reversal shift, he would have Form retained as the relevant dimension but with Triangle as the positive cue. In a nonreversal shift, the relevant dimension would be Color and either red or green would be the positive cue depending on his assignment. These arrangements were used in each of the Training conditions of both experiments.